

Optimization and Follow-up of Semicoherent Searches for Continuous Gravitational Waves

Von der Fakultät für Mathematik und Physik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades

Doktor der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation
von

Dipl.-Phys. Miroslav Shaltev

geboren am 1. Mai 1981, in Sofia

2013

Referent: Prof. Bruce Allen

Korreferent: Prof. Andrzej Krolak

Tag der Promotion: 25/06/2013

Abstract

Direct detection of gravitational-wave emission is expected not only to further increase our confidence in the validity of Einstein's General Relativity theory, but generally enrich our understanding of the universe. The subject of this thesis is the improvement of the data analysis techniques in the search for previously-unknown objects, such as rapidly spinning neutron stars, which are possible emitters of continuous gravitational waves.

Coherent wide parameter-space searches for continuous gravitational waves are typically limited in sensitivity by their prohibitive computing cost. Therefore semicoherent methods, such as StackSlide can often achieve a better sensitivity. An analytical method for finding optimal StackSlide parameters at fixed computing cost is presented. This method assumes ideal conditions of gap-less data with Gaussian stationary noise. The obtained solution suggests either the usage of all of the available data, or a finite optimal observation time. Various practical examples for the application of this optimization framework are given, illustrating the potential gains in sensitivity compared to previous searches.

The analytical framework to find optimal semicoherent search parameters is then extended to a numerical optimization method, in order to take into account the possible gaps in real detector data and noise level changes. This results in the practical implementation of a data selection procedure. For comparison with previously obtained results the numerical optimization method is first applied to ideal data, and then the ideal conditions are gradually relaxed. Finally, the performance of this method is illustrated on real detector data.

The interesting parameter-space regions selected in a wide parameter-space semicoherent search need fully coherent follow-up investigation in order to confirm a possible gravitational wave emitter. Two different methods to compute the minimal required integration time of a fully coherent follow-up search are derived and compared.

A general framework for the transition from a semicoherent to a fully coherent search using all the data in two stages is developed. The practical implementation of this two-stage follow-up scheme is suitable for systematic follow-up of candidates at feasible computational cost.

The two-stage follow-up procedure is used to investigate candidates from the Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data. All hardware injections identified by the Einstein@Home search are recovered and another nine outliers of unknown origin are followed-up. No significant gravitational wave events are found.

Keywords: continuous gravitational wave detection, follow-up searches, data analysis.

Kurzfassung

Die direkte Beobachtung von Gravitationswellen wird nicht nur die Zuversicht in der Allgemeine Relativitätstheorie steigern, sondern auch im allgemein das Wissen über das Universum weiter bereichern. Das Thema der vorliegenden Dissertation ist die Weiterentwicklung der Methoden für die Suche nach unbekannten Objekten, zum Beispiel schnell-rotierende Neutronensterne, die kontinuierliche Gravitationswellen abstrahlen können.

Die Sensitivität der kohärente Suche nach kontinuierlichen Gravitationswellen ist beschränkt durch die hohe erforderliche Rechenleistung. Entsprechend erreichen inkohärente Methoden, z.B. StackSlide, oft bessere Sensitivität. Zuerst wird eine analytische Methode präsentiert, um die optimalen Parameter der StackSlide-Suche bei konstanter Rechenleistung zu ermitteln. Dabei werden ideale Bedingungen angenommen, nämlich kontinuierliche Daten und stationäres gaußsches Rauschen. Es wird gezeigt, dass entweder alle Daten benutzt werden sollen, oder es gibt eine optimale Beobachtungszeit. Verschiedene Beispiele von praktischer Bedeutung und die mögliche Verbesserung der Sensitivität werden diskutiert.

Die analytische Methode wird dann mit einer numerischen Optimierung erweitert um den diskreten Charakter der reellen Daten, wie auch die Fluktuation des Rauschens, in Betracht zu ziehen. Die praktische Umsetzung der Methode führt zu einem Verfahren zur optimalen Datenauswahl. Die numerische Optimierung wird zunächst unter idealen Bedingungen eingesetzt und mit der analytischen Methode verglichen. Danach werden die Anforderungen an der Daten gelockert und zum Schluss werden reelle Detektordaten benutzt.

Die Kandidaten einer inkohärenten Suche benötigen eine kohärente Nachuntersuchung. Zwei verschiedene Methoden für die Berechnung der minimalen erforderlichen Beobachtungszeit werden präsentiert.

Es wird ein allgemeines Verfahren zum Übergang von inkohärenten zu kohärenten Suchen entwickelt. Durch die praktische Implementierung einer zwei-stufigen Version des Verfahrens wird gezeigt, dass eine systematische Nachuntersuchung von inkohärenten Kandidaten mit kohärenten Methoden mit realistischen Rechenleistungen realisierbar ist.

Die zwei-stufige Nachuntersuchungsmethode wird an Kandidaten aus der Einstein@Home Suche in LIGO S5 Daten angewandt. Die hardwaresimulierten Signale, die durch die inkohärente Suche identifiziert wurden, werden vollständig bestätigt. Neun weitere Kandidaten werden mit dieser Methode untersucht, keiner davon wird als eine reelle Gravitationswelle nachgewiesen.

Schlagwörter: kontinuierliche Gravitationswellen, Nachuntersuchungsmethoden, Daten Auswertung.

Contents

Abstract	i
Kurzfassung	iii
Contents	v
List of Figures	ix
List of Tables	x
Synopsis	xii
I. Introduction	1
1. Gravitational waves	3
1.1. Linearized gravity and wave generation	3
1.2. Gravitational-wave detectors	6
1.2.1. Resonant-mass detectors	6
1.2.2. Interferometric detectors	7
1.2.3. The detector response to a gravitational wave and the noise output	11
1.3. Sources of gravitational waves	12
1.3.1. Stochastic Gravitational-Wave Background and Gravitational-Wave Bursts	12
1.3.2. Continuous Gravitational Waves	13
2. Searches for continuous gravitational waves	16
2.1. Signal model	16
2.2. Detection statistic	20
2.2.1. Hypothesis testing	20
2.2.2. The $\mathcal{F}$ -statistic	21
2.3. $\mathcal{F}$ -statistic based searches	24
2.3.1. Signal-to-noise ratio	25
2.3.2. Wide parameter-space searches	29
2.3.3. The choice of spin-down search band	31

3. Optimization by Direct Search	35
3.1. Stochastic optimization	35
3.2. Mesh Adaptive Direct Search	35
3.2.1. The <i>search</i> step	36
3.2.2. The <i>poll</i> step	37
3.2.3. NOMAD	40
 II. Semicohherent Search Optimization	 43
4. Optimal StackSlide method at fixed computing cost	46
4.1. Maximizing probability of a CW detection	48
4.2. Properties of a single-stage StackSlide search	49
4.2.1. The StackSlide search method	50
4.2.2. Mismatch and metric	50
4.2.3. Sensitivity estimate	55
4.2.4. Template counting	60
4.2.5. Computing-cost model	61
4.3. Maximizing sensitivity at fixed computing cost	64
4.3.1. Special case (i): Fully coherent search	65
4.3.2. Special case (ii): Computing cost dominated by one contribution	65
4.3.3. General optimality conditions	67
4.4. Examples of practical application	70
4.4.1. Directed searches for isolated neutron stars	71
4.4.2. All-sky CW search using Einstein@Home	75
4.4.3. All-sky search examples from CGK	76
4.4.4. CWs from binary neutron stars	77
4.5. Discussion	78
 5. Optimal StackSlide method at fixed computing cost using real data	 80
5.1. Threshold SNR, sensitivity function and computing cost	81
5.1.1. Threshold SNR	81
5.1.2. Sensitivity function	82
5.1.3. Computing cost	84
5.1.4. Templates counting	85
5.1.5. The choice of spindown parameter space	86
5.2. Numerical optimization procedure for a semicoherent StackSlide search with a fixed frequency band	86
5.2.1. Definition of the optimization problem	86
5.2.2. Data selection	87
5.3. Examples of practical application	88
5.3.1. Directed search using simulated data	89
5.3.2. Directed search using real data	94

5.4. Discussion	97
III. Fully coherent Follow-up Studies	99
6. Minimal required observation time	102
6.1. Properties of $\mathcal{F}$ -statistic searches	102
6.1.1. Fully coherent search	103
6.1.2. Semicoherent search	103
6.1.3. Template counting	104
6.1.4. Computing cost	105
6.2. Minimal required observation time	106
6.2.1. Method 1	106
6.2.2. Method 2	107
6.3. Method comparison	108
6.3.1. Numerical predictions	108
6.3.2. Monte Carlo results	109
6.4. Discussion	111
7. Fully coherent follow-up of continuous gravitational-wave candidates	112
7.1. Continuous Gravitational Waves	113
7.1.1. Detection statistic	113
7.1.2. Mismatch and Fisher matrix	114
7.1.3. Computing cost	116
7.2. Coherent follow-up of semicoherent candidates	116
7.2.1. Basic two-stage search strategy	116
7.2.2. Classification of zoom outcomes	120
7.2.3. Grid-based computing cost of the zoom stage	121
7.2.4. Mesh Adaptive Direct Search (MADS)	122
7.2.5. MADS-based follow-up algorithm	123
7.2.6. MADS follow-up computing cost	125
7.2.7. False-alarm and detection probability	125
7.3. Monte Carlo studies	126
7.3.1. Follow-up of candidates from a directed search	127
7.3.2. Follow-up of candidates from an all-sky search	129
7.4. Discussion	131
8. Fully coherent follow-up of candidates from the Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data	132
8.1. The LIGO S5 all-sky search with Einstein@Home	133
8.2. Follow-up method	135
8.2.1. The two-stage follow-up	135
8.2.2. A pre-refinement with the semicoherent Hough-transform search	136

Contents

8.2.3. The number of search dimensions	136
8.3. Monte Carlo studies	137
8.3.1. Efficiency of the follow-up pipeline	140
8.3.2. Impact of second-order spindown on the first-order spindown follow-up	140
8.4. Recovery of hardware injections	142
8.5. Follow-up of loud S5R5 Einstein@Home candidates	143
8.6. Discussion	145
References	146
Acknowledgements	158
Curriculum Vitae	160

List of Figures

1.1.	The effect of a gravitational wave passing through a circle of free particles.	5
1.2.	A resonant bar and a spherical detector.	7
1.3.	Schematic layout of an interferometric gravitational-wave detector.	8
1.4.	Aerial view of the first generation of large-scale gravitational-wave detectors.	10
1.5.	Strain sensitivity during S5 and VSR1.	13
2.1.	Signal model.	17
2.2.	Example plot of χ^2 pdf.	26
2.3.	Spindown search volume.	34
3.1.	The 2D Rosenbrock function and a quadratic model of it.	38
3.2.	Example of MADS frames.	41
3.3.	MADS optimization of the 2D Rosenbrock function.	42
4.1.	Mismatch triangle.	54
4.2.	N -scaling coefficient w .	59
4.3.	Numerical optimal solution.	73
5.1.	Ratio of threshold SNR computed with the WSG and KWS approximation.	83
5.2.	Schematic representation of the data selection.	87
5.3.	Numerical optimization using ideal data.	91
5.4.	Numerical optimization using data with constant noise-floor and gaps.	92
5.5.	Numerical optimization using simulated data with gaps.	93
5.6.	Numerical optimization using real data.	95
5.7.	Numerical optimization using real data for an Einstein@Home search.	96
6.1.	Minimal required coherent integration time.	110
7.1.	Two-dimensional search grid in $\{f, \dot{f}\}$ space.	117
7.2.	Two-dimensional example: Fisher ellipse.	119
7.3.	MADS-based search algorithm with four passes.	124
7.4.	MC study of two-stage follow-up of candidates from a directed search.	128
7.5.	MC study of two-stage follow-up of candidates from an all-sky search.	130
8.1.	Zoom stage on Gaussian noise.	137
8.2.	Monte Carlo study of the efficiency of the S5R5 follow-up pipeline.	138
8.3.	MC study of the impact of second-order spindown.	141

List of Tables

4.1. Overview of symbols and notation for the analytical optimization.	64
4.2. Einstein@Home example setups 'S5GC1' and 'S6Bucket'.	75
4.3. CGK example search setups.	77
5.1. Detector data used to test the numerical optimization under real conditions.	94
5.2. Summary of the results of the numerical optimization.	97
6.1. Joint probability p_J and corresponding required h standard deviations. .	108
7.1. Algorithm parameters for follow-up of candidates from directed searches.	127
7.2. Follow-up algorithm parameters for full parameter space searches.	129
8.1. Follow-up algorithm parameters for the refinement and zoom stage. . . .	140
8.2. Loudest candidates associated with the hardware injections.	144
8.3. Follow-up results for the nine candidates from the S5R5 search.	145

Synopsis

The thesis is divided into three parts:

- Part I - *Introduction*. In Chap. 1 we discuss the generation of gravitational waves (GW) in linearized gravity, GW detectors and GW sources. Chap. 2 is dedicated to the searches for continuous gravitational waves, in particular the signal model and detection statistics with emphasis on $\mathcal{F}$ -statistic based searches. This part is closed with a short Chap. 3, describing the Mesh Adaptive Direct Search class of algorithms for constraint optimization, which will be applied in later chapters.
- Part II - *Semicoherent Search Optimization*. This part is focused on studying the optimal semicoherent search parameters required to maximize the sensitivity of a search at fixed computing cost. Chap. 4 provides analytical treatment of the problem using the StackSlide semicoherent search under ideal conditions. Semicoherent search optimization using real data is considered in Chap. 5.
- Part III - *Fully-coherent Follow-up Studies*. The final part considers the treatment of outliers of semicoherent searches. In Chap. 6, we derive two different methods to compute the minimal required observation time in order to distinguish gravitational-wave candidates from Gaussian noise at a given false-alarm and false-dismissal probability. Chap. 7 introduces a follow-up strategy suitable for systematic follow-up of candidates using all of the available data. An application of this method is presented in Chap. 8, where we first recover the hardware injections identified in a recent Einstein@Home search for continuous gravitational waves, and then discard other loud outliers as consistent with Gaussian noise.

The introductory part of this work is based on references [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13] and references therein, although additional publications are referenced where appropriate. Part II and III are based on my published papers, e.g., Chaps. 4, 6 and 7, and publications in preparation, e.g., Chaps. 5 and 8. This work has been done under the supervision and in collaboration with Dr. Reinhard Prix.

List of publications

Refereed publications

1. M. Shaltev and R. Prix, “Fully coherent follow-up of continuous gravitational-wave candidates”, *Phys. Rev. D*, vol. 87, p. 084057, Apr 2013
2. M. Shaltev, “Coherent follow-up of continuous gravitational-wave candidates: minimal required observation time” *Journal of Physics: Conference Series*, vol. 363, no. 1, p. 012043, 2012
3. R. Prix and M. Shaltev, “Search for continuous gravitational waves: Optimal stackslide method at fixed computing cost” *Phys. Rev. D*, vol. 85, p. 084010, Apr 2012

Publications in preparation

1. M. Shaltev and R. Prix, “Optimal StackSlide method at fixed computing cost using real data”
2. M. Shaltev, P. Leaci, M. A. Papa and R. Prix, “Fully coherent follow-up of loud candidates from the Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data”

Contributed to

1. J. Aasi *et al.*, “Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data” *Phys. Rev. D*, vol. 87, p. 042001, Feb 2013

Part I.

Introduction

1. Gravitational waves

Around 1936 Einstein arrived at a remarkable conclusion about the existence of gravitational-wave radiation, a consequence of his own theory of General Relativity (GR), namely [14]:

“Together with a young collaborator¹, I arrived at the interesting result that gravitational waves do not exist, though they had been assumed a certainty to the first approximation.”

However, Einstein accepted later that this conclusion had been erroneous, as pointed out by Robertson and Infeld. Some 40 years later, in 1975, the discovery of the first binary pulsar (PSR1913+16) by Hulse and Taylor [15] provided the first *indirect evidence for the existence* of gravitational waves.

This chapter is organized as follows. In Sec. 1.1 we discuss linearized gravity and wave generation, and in Sec. 1.2 we discuss gravitational-wave detectors. Possible sources, in particular of continuous gravitational waves, are discussed in Sec. 1.3.

1.1. Linearized gravity and wave generation

Following Schutz [1], gravity and its action on matter in GR is described based on the idea of a curved manifold with a metric tensor $\mathbf{g}$. The law how the source of the gravitational field, namely the stress-energy tensor $\mathbf{T}$, determines the metric is given by Einstein’s equation, which in component notation and geometrized units ($c = 1$, $G = 1$) are

$$G^{\mu\nu} = 8\pi T^{\mu\nu} , \quad (1.1)$$

where $G^{\mu\nu}$ is the Einstein tensor

$$G^{\mu\nu} \equiv R^{\mu\nu} - \frac{1}{2}g^{\mu\nu}R , \quad (1.2)$$

with $R^{\mu\nu}$ the symmetric Ricci tensor and $R = g^{\mu\nu}R_{\mu\nu}$ the Ricci scalar². The Ricci tensor is obtained by contraction of the Riemann curvature tensor

$$R^{\mu}_{\nu\alpha\beta} = \partial_{\alpha}\Gamma^{\mu}_{\nu\beta} - \partial_{\beta}\Gamma^{\mu}_{\nu\alpha} + \Gamma^{\mu}_{\sigma\alpha}\Gamma^{\sigma}_{\nu\beta} - \Gamma^{\mu}_{\sigma\beta}\Gamma^{\sigma}_{\nu\alpha} , \quad (1.3)$$

¹Nathan Rosen

²We use the Einstein summation convention summing over repeated indices. We raise and lower indices using the metric.

on the first and third indices, i.e., $R_{\mu\nu} = R^\alpha_{\mu\alpha\nu}$, and

$$\Gamma^\gamma_{\nu\alpha} = \frac{1}{2}g^{\gamma\mu}(\partial_\alpha g_{\mu\nu} + \partial_\nu g_{\mu\alpha} - \partial_\mu g_{\nu\alpha}) \quad (1.4)$$

are the Christoffel symbols.

In the limit of a weak gravitational field the metric tensor $g_{\mu\nu}$ can be expressed as a small perturbation $|h_{\mu\nu}| \ll 1$ of the Minkowski metric $\eta_{\mu\nu}$, i.e.,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} . \quad (1.5)$$

Keeping terms linear in $h_{\mu\nu}$, with its trace reverse

$$\bar{h}^{\mu\nu} \equiv h^{\mu\nu} - \frac{1}{2}\eta^{\mu\nu}h^\alpha_\alpha , \quad (1.6)$$

and using the Lorentz gauge

$$\partial_\nu \bar{h}^{\mu\nu} = 0 , \quad (1.7)$$

the Einstein tensor can be written in the form

$$G^{\mu\nu} = -\frac{1}{2}\square\bar{h}^{\mu\nu} , \quad (1.8)$$

where $\square \equiv \eta^{\mu\nu}\partial_\mu\partial_\nu$ is the wave operator. Thus the *linearized* formulation of (1.1) is

$$\square\bar{h}^{\mu\nu} = -16\pi T^{\mu\nu} , \quad (1.9)$$

Far outside the source of the field with $T^{\mu\nu} = 0$ we therefore obtain the wave equation

$$\square\bar{h}^{\mu\nu} = 0 , \quad (1.10)$$

with solutions of the form

$$\bar{h}^{\mu\nu} = H^{\mu\nu} e^{i(\omega t - \vec{k}\vec{x})} , \quad (1.11)$$

where $\vec{k}$ is the wave vector, ω is the angular frequency of the wave and $\mathbf{H}$ is a tensor. The solution (1.11) is a plane wave. In the *transverse-traceless* (TT) gauge we have

$$\bar{h}^{TT}_{\mu\nu} = h^{TT}_{\mu\nu} . \quad (1.12)$$

In a frame where the wave is traveling in the z -direction, $H^{TT}_{\mu\nu}$ takes the form

$$H^{TT}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} , \quad (1.13)$$

where h_+ and $h_\times$ are the polarization amplitudes. Consider two particles at rest, one

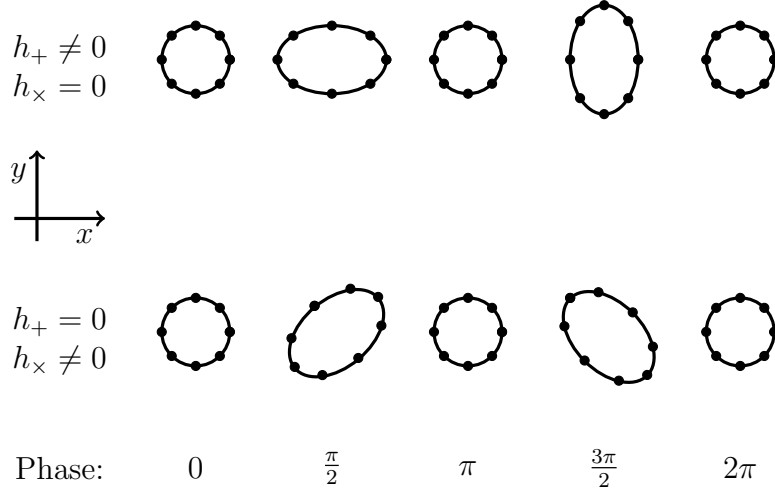


Figure 1.1.: The effect of a gravitational wave passing in z -direction through a circle of free particles.

at the origin and another one at $x = l$, $y = z = 0$. Computation of the proper distance yields [10]

$$\begin{aligned}
 \Delta l &= \int_0^l dx \sqrt{g_{xx}} \\
 &= \int_0^l dx \sqrt{1 + h_{xx}^{TT}} \\
 &\approx \int_0^l dx \left(1 + \frac{1}{2} h_{xx}^{TT}\right) \\
 &\approx \left(1 + \frac{1}{2} h_{xx}^{TT}\right) l .
 \end{aligned} \tag{1.14}$$

This means that the larger the separation between the test particles, the larger the effect of a passing gravitational wave. Furthermore, the effect is proportional to the perturbation, h_{xx}^{TT} , and thus, small. The effect of a gravitational wave passing in z -direction through a circle of free particles is schematically depicted in Fig. 1.1.

In the limit where the gravitational-wave length is much smaller compared to the distance to the source, but much larger than its size, the perturbation h can be computed using the “quadrupole formula” [5]

$$h_{ij}^{TT}(t) \simeq \frac{2G}{c^4 r} \ddot{I}_{ij}^{TT}(t - r/c) , \tag{1.15}$$

where

$$I^{ij}(t) = \int d^3x x^i x^j \rho(t, x) \tag{1.16}$$

is the quadrupole tensor with mass density $\rho(t, x)$ and

$$I_{ij}^{TT} = P_{ik} I^{kl} P_{lj} - \frac{1}{2} P_{ij} P_{kl} I^{kl} \quad (1.17)$$

is its projection using the transverse projection operator

$$P_{ij} = \delta_{ij} - \hat{n}_i \hat{n}_j, \quad (1.18)$$

with $\hat{n}^i = x^i/r$ the unit vector in the propagation direction. Using the quadrupole tensor, the gravitational-wave luminosity is expressed as [5]

$$L = \frac{1}{5} \frac{G}{c^5} \left\langle \ddot{I}_{ij} \ddot{I}^{ij} \right\rangle, \quad (1.19)$$

where $\langle . \rangle$ denotes the time average over several periods.

1.2. Gravitational-wave detectors

The pioneering work of Joseph Weber in the 1960s at Argonne National Laboratory and at the University of Maryland [16] is considered as the beginning of experimental gravitational-wave physics, further driven by Ronald Drever and Rainer Weiss in the 1990s.

There are mainly two conceptually different types of devices designed for the detection of gravitational radiation. The earlier technology of Resonant-mass detectors, also known as *bar* detectors, rely on the excitement of acoustic modes of oscillation of the bar due to a passing gravitational wave. The second type, interferometric detectors, use laser interferometry to measure the change of the optical path length produced by gravitational radiation [5].

1.2.1. Resonant-mass detectors

The past, present, and future of this type of detector is comprehensively reviewed by Aguiar [17]. The bar detectors were most successful in the 90's with their third generation, where the bar network of 5 detectors, namely Allegro (Baton Rouge, Louisiana, USA), Auriga (Padova, Italy) Fig. 1.2(a), Explorer (CERN, Geneva, Switzerland), Nautilus (Roma, Italy) and Niobe (Perth, Australia), has been in continuous operation mode. With advancements of interferometric detectors, bar detectors have been outperformed in terms of sensitivity and frequency range. There is still some effort to overcome these problems by using spherical (MiniGRAIL Fig. 1.2(b), Mario Schenberg) instead of cylindrical detectors and so-called “dual bar” detectors [18][19][20].

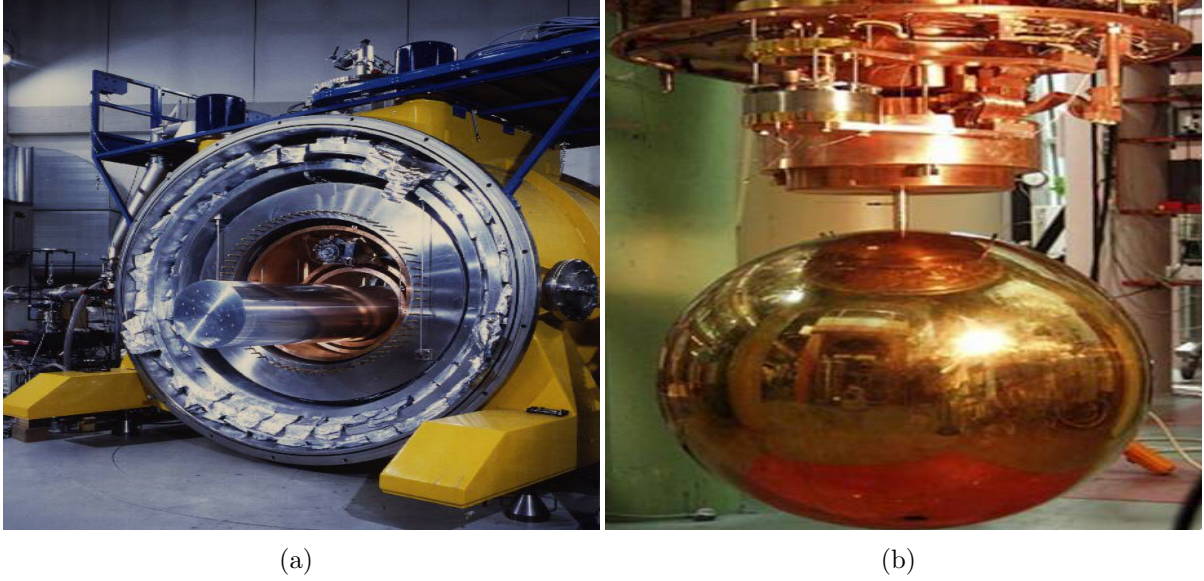


Figure 1.2.: (a) A resonant bar (Auriga). (b) A spherical detector (MiniGRAIL).

1.2.2. Interferometric detectors

A simple interferometric gravitational-wave detector can be realized with a *Michelson* interferometer. In such a device, a beam of laser light is sent through a beam splitter in two directions, the arms of the interferometer. At the end of the arms, the light is reflected back by the mirrors. The suspended mirrors represent freely-falling test masses in the plane of the interferometer. If the arm lengths are equal, up to an integer number of the wavelength, then the light travels back to the laser. Otherwise a fraction of it can be detected at the output port. Thus, if the device is set up, such that light is not transmitted in the absence of gravitational radiation, a passing gravitational wave could be detected, by observing the light at the output port when one of the arms is squeezed, while the other is stretched by the passing gravitational wave. The strain, $h(t)$, caused by the wave is proportional to the ratio of the change in the distance in the arm length Δl to the arm length l [5]

$$h(t) \approx \frac{\Delta l}{l} . \quad (1.20)$$

In an interferometric GW detector the strain sensitivity is constrained by the following noise sources (though this is a non-exhaustive list of noise sources):

- *Shot Noise* - the measurement of the output power consist effectively in counting of the number of photons at the output photodetector and improves with higher laser power.
- *Radiation Pressure Noise* - is introduced by the fluctuating number of photons

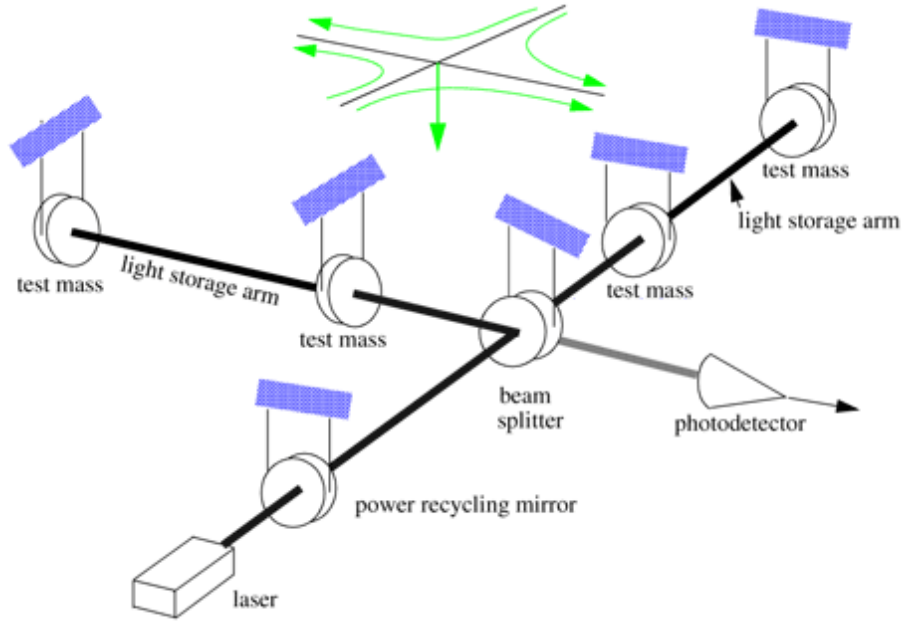


Figure 1.3.: Schematic layout of an interferometric gravitational-wave detector [21], displaying power-recycling and Fabry-Perot cavities.

reflecting off the mirrors, thus acting with a stochastic force, which shakes the test masses.

- *Seismic Noise* - the Earth's ground is not static but in continuous vibration due to earthquakes, human activities and local phenomena like winds and a micro-seismic background. For ground-based detectors, attenuation filters, e.g., masses and springs, and / or a set of pendulums, can reduce this type of noise to levels interesting for gravitational-wave detection for frequencies above 10 Hz [3].
- *Thermal Noise* - appears in the suspensions and the test masses and depends on the temperature, e.g., the Brownian motion of the coatings of the mirrors.
- *Gravity Gradient Noise* - seismic surface waves yield fluctuation of the Earth's density, which gravitationally couples with the mirrors. This effect can be actively attenuated or as alternative the detector can be placed underground [22][23].

Some examples of upgrades to simple Michelson interferometers to improve the sensitivity are[22]:

- *Fabry-Perot cavity* - adding a partially transmissive mirror along an optical axis tuned such that light power builds-up inside the cavity effectively increases the arm length; see Fig. 1.3.

- *Power Recycling* - light that travels back to the laser and would normally be lost can be re-injected to increase the amount of laser power circulating in the interferometer; see Fig. 1.3.
- *Signal Recycling* - sidebands created on the light due to a gravitational-wave signal will appear at the output port and can be reflected back into the detector for further amplification.
- *Squeezed Light* - injection of squeezed states of light in the output of the interferometer can improve shot noise limited sensitivity [24].

The initial large-scale detectors

The initial large-scale interferometric detectors include the Laser Interferometric Gravitational-wave Observatory (LIGO) in Hanford and Livingston (USA), the VIRGO detector in Pisa (Italy), the GEO600 detector near Hannover (Germany) and the TAMA300 detector in Tokyo (Japan). Aerial photographs of the first four detectors, which make up the detector network of the LSC-Virgo (LVC) collaboration, are shown in Fig. 1.4.

The TAMA300 power-recycled Michelson interferometer was the first large-scale detector in operation with 300 m long arms.

LIGO Hanford (LHO) operated two power-recycled Michelson interferometers with Fabry-Perot 4 km (H1) and 2 km (H2) long arm cavities. Similarly, LIGO Livingston (LLO) operated one power-recycled Michelson interferometer with 4 km arm cavities. The Virgo detector is also a power-recycled Michelson interferometer with 3 km long arm cavities. After the fifth Science Run (S5) of the LIGO detectors and the first Virgo Scientific Run (VSR1) of the Virgo detector, the instruments were upgraded for the S6 and VSR2/3 joint run. Presently, the initial detectors have been decommissioned in order to install the next generation of interferometric detectors: Advanced LIGO and Virgo

The GEO600 detector is a dual-recycled Michelson interferometer with 600 m long arms. While less sensitive compared to the other three detectors of the LVC, it has been used to develop technologies for the advanced detectors. Further, this is the only large scale detector remaining in operation until the advanced detectors come online.

Future gravitational-wave detectors

The second generation of ground based detectors is currently being installed. Using higher laser power, more massive mirrors, and better seismic noise attenuation, these advanced detectors are expected to have an improved strain sensitivity by approximately an order of magnitude [25]. In addition, the world-wide network of detectors will likely be extended with one of the Hanford LIGO detectors being installed in India and another large-scale interferometer, Kagra, in Japan [26] whose novel features include cryogenically cooled mirrors installed underground.

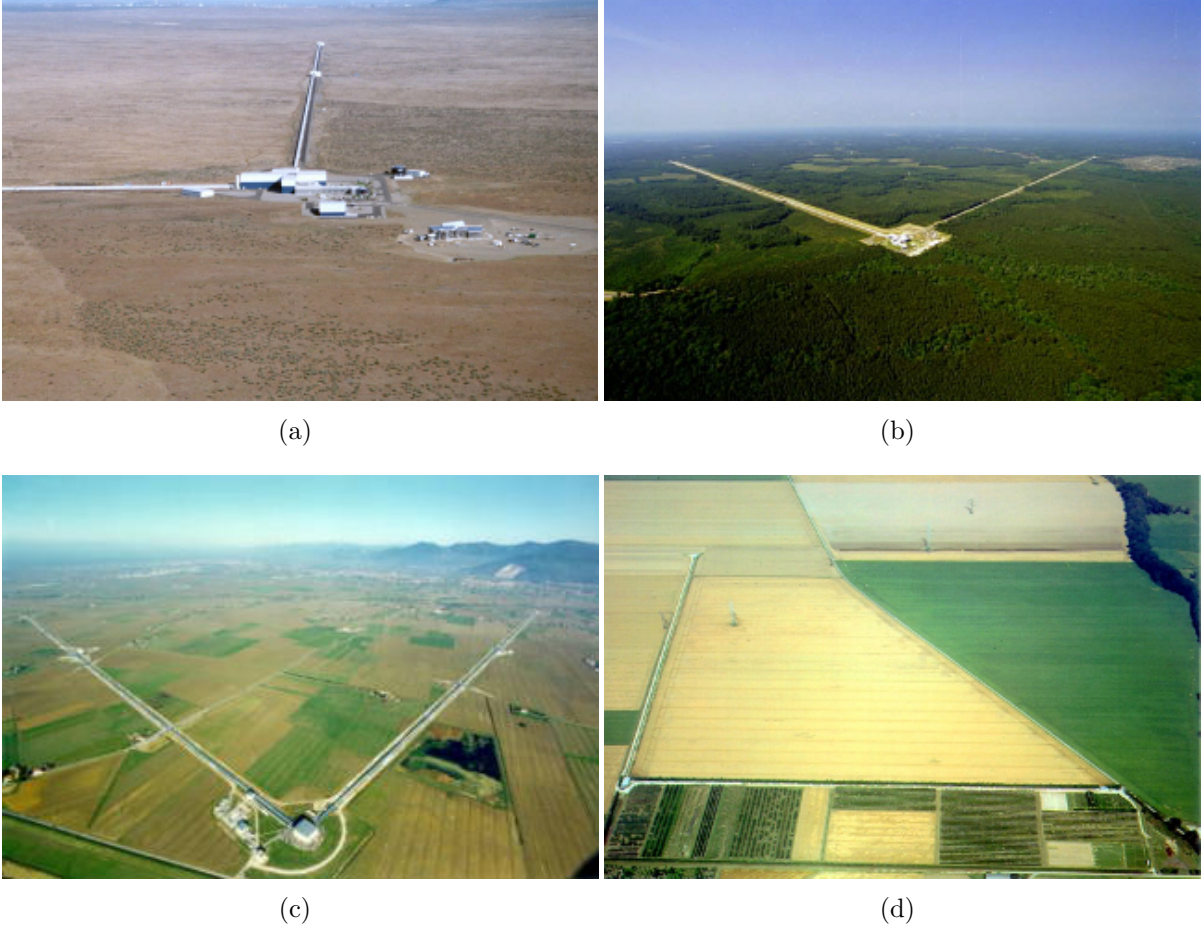


Figure 1.4.: Aerial view of the first generation of large-scale gravitational-wave detectors: (a) LIGO Hanford, (b) LIGO Livingston, (c) Virgo, and (d) GEO600.

The design study of a third generation detector, the Einstein Telescope (ET), has been finished, proposing a large gravitational-wave observatory underground [27]. The detector consist of 3 interferometers with 10-km-long arms nested in a triangle shape. The expected sensitivity of ET is an order of magnitude better than the expected sensitivity of the second generation of detectors.

Finally, the space-born observatory, LISA, still periodically redesigned due to mostly non-scientific constraints, is expected to bring valuable knowledge about sources emitting in the low-frequency range $[10^{-4} \text{ Hz} - 10^{-1} \text{ Hz}]$ of the gravitational-wave spectrum.

1.2.3. The detector response to a gravitational wave and the noise output

Denoting the two arm lengths of an interferometric detector with l_1 , resp. l_2 , the measured strain of the detector is defined as:

$$h(t) \equiv \frac{\Delta l_1}{l_1} - \frac{\Delta l_2}{l_2} . \quad (1.21)$$

Assuming that the arm lengths of the detector are much shorter than the gravitational wavelength, the measured strain from the above equation can be written in terms of the gravitational-wave polarizations as [28]:

$$h(t) = F_+(t)h_+(t) + F_\times(t)h_\times(t) , \quad (1.22)$$

where the functions $F_+(t)$ and $F_\times(t)$ are the so called antenna-patterns. These functions depend on the source position and polarization, as well as on the orientation of the detector. More details will be given in Chap. 2. For now it is sufficient to assume, that the measurable strain is a small quantity, i.e., the signal is buried deep in the noise of the detector.

Independent of the type of the detector, the collected data is a time series $x(t)$. We assume that the data is noise dominated $x(t) \approx n(t)$ and represented by a stationary random process, i.e., we have a constant zero mean

$$\langle n \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} dt n(t) = 0 , \quad (1.23)$$

and the covariance

$$C(n(t_1), n(t_2)) \equiv E[(n(t_1) - E[n(t_1)])(n(t_2) - E[n(t_2)])] , \quad (1.24)$$

depends only on the time difference $t_1 - t_2$, where $E[x] \equiv \int dx x p(x)$ is the expectation value. It is also useful to define the Fourier transform of a time series $x(t)$, which yields the frequency series $\tilde{x}(f)$, i.e., [5]

$$\tilde{x}(f) = \int_{-\infty}^{\infty} dt x(t) e^{-2\pi i f t} , \quad (1.25)$$

and the inverse Fourier transform

$$x(t) = \int_{-\infty}^{\infty} df \tilde{x}(f) e^{2\pi i f t} \quad (1.26)$$

Using Parseval's theorem,

$$\int_{-\infty}^{\infty} dt |x(t)|^2 = \int_{-\infty}^{\infty} df |\tilde{x}(f)|^2 , \quad (1.27)$$

the power of the noise is [5]

$$\begin{aligned}
\langle n^2 \rangle &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} dt n_T^2(t) \\
&= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-\infty}^{\infty} df |\tilde{n}_T(f)|^2 \\
&= \lim_{T \rightarrow \infty} \frac{2}{T} \int_0^{\infty} df |\tilde{n}_T(f)|^2 \\
&= \int_0^{\infty} df S_n(f) ,
\end{aligned} \tag{1.28}$$

where $n_T(t)$ is the windowed data

$$n_T(t) = n(t) , \tag{1.29}$$

if $t \in (-T/2, T/2)$ and otherwise 0. The quantity $S_n(f)$ is the (single-sided) *power spectral density* [5]

$$S_n(f) \equiv \lim_{T \rightarrow \infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} dt n(t) e^{-2\pi i f t} \right|^2 . \tag{1.30}$$

With this, we characterize a detector using the strain spectral sensitivity, or spectral amplitude, $\sqrt{S_n(f)}$, which has dimension $1/\sqrt{\text{Hz}}$. For example, the sensitivity curves of the LIGO, VIRGO and GEO600 detectors during the S5 and VSR1 Science Runs are plotted in Fig. 1.5.

1.3. Sources of gravitational waves

Depending on the time evolution of the gravitational waveform, we distinguish three major classes of astrophysical gravitational wave sources: bursts, stochastic and periodic.

1.3.1. Stochastic Gravitational-Wave Background and Gravitational-Wave Bursts

The stochastic gravitational-wave background is a signal originating from the contribution of a large number of individual sources that cannot be distinguished from each other. This is commonly referred to as source confusion. These signals could be of cosmological or astrophysical origin. Although detection of cosmological sources with foreseeable technologies is unlikely, observation of the astrophysical stochastic background is expected to happen with a space-born gravitational-wave observatory, e.g., LISA [5].

Gravitational-wave bursts are short-lived signals (compared to the observation time) produced in some catastrophic astrophysical event, e.g., gravitational collapse, pulsar glitches, or another highly energetic event. These signals are either unmodeled or modeled. Unmodeled signals have to be searched without any other external information

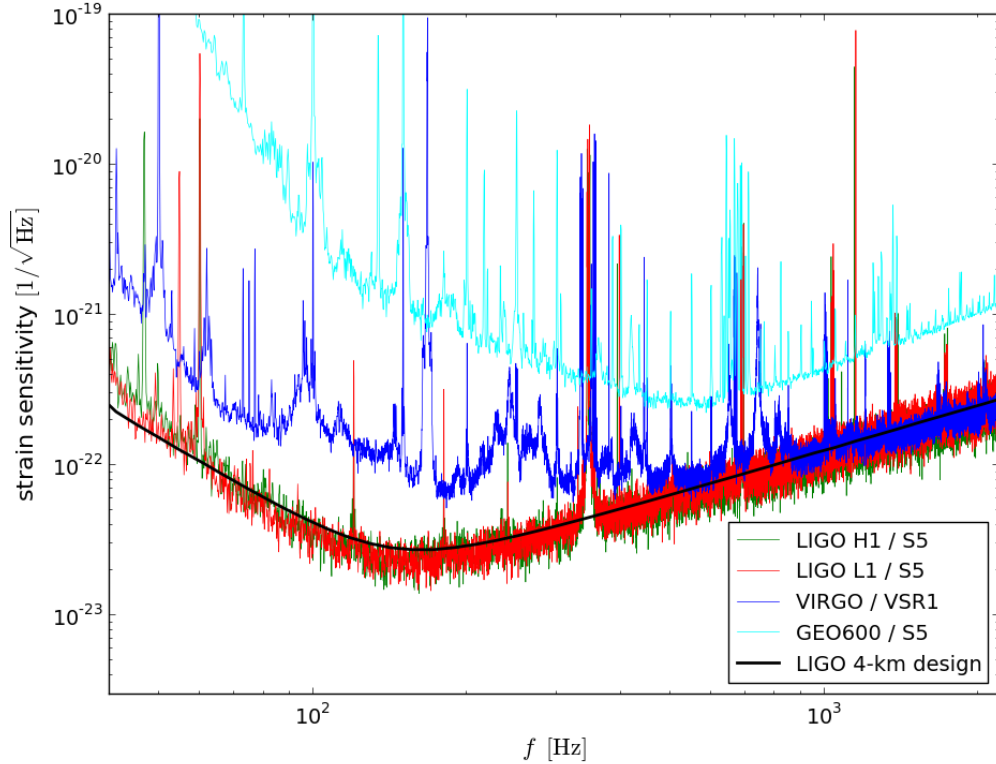


Figure 1.5.: LIGO, VIRGO and GEO600 strain sensitivity during the S5 / VSR1 Science Runs. Note that the LIGO detectors reached their design sensitivity.

[29], or by using external triggers, e.g., gamma-rays [30] or neutrino emission [31] to constrain the parameter space. The prototypical example for modeled burst sources are compact binary coalescences. The coalescence of compact binary stars, e.g., white dwarfs, neutron stars or black holes, is expected to be the most promising source for detection with future detectors [5]. The expected signal is sinusoidal with increasing frequency and amplitude, a chirp.

1.3.2. Continuous Gravitational Waves

A continuous gravitational-wave signal is a long-lived signal, emitted, for example, by rapidly spinning neutron stars. The emission of GWs from such objects is primarily driven by non-axisymmetric distortion, non-axisymmetric instability or free precession [32].

Non-axisymmetric distortion

A non-axisymmetric spinning neutron star at distance d emits gravitational waves at a frequency f , that equals twice the rotational frequency ν , i.e., $f = 2\nu$, with a character-

istic amplitude

$$h_0 = \frac{16\pi^2 G I_3 \nu^2}{c^4 d} \epsilon, \quad (1.31)$$

where G is the gravitational constant, I_3 is the moment of inertia about the rotational axis, and

$$\epsilon = \frac{|I_1 - I_2|}{I_3} \quad (1.32)$$

is the ellipticity. I_1 and I_2 are the other moments of inertia about principle axis. The possible values of the ellipticity are highly uncertain. An estimate of the maximal ellipticity is given by [32]

$$\epsilon = 5 \times 10^{-7} \left(\frac{\sigma}{10^{-2}} \right), \quad (1.33)$$

where σ is the breaking strain of the solid crust. For an ellipticity $\epsilon \sim 10^{-6}$ [33], using Eq. (1.31) we obtain

$$h_0 \approx 1 \times 10^{-26} \left(\frac{\epsilon}{10^{-6}} \right) \left(\frac{I}{I_{38}} \right) \left(\frac{1 \text{ kpc}}{d} \right) \left(\frac{f}{100 \text{ Hz}} \right)^2, \quad (1.34)$$

where $I_{38} = 10^{38} \text{ kg m}^2$ is the typical moment of inertia for a neutron star [34].

Non-axisymmetric instabilities

Rapidly rotating neutron stars can undergo various non-axisymmetric instabilities at birth or during accretion. Hydrodynamics and gravity could drive dynamical instability, if the star is very rapidly rotating, i.e., if the ratio of the kinetic energy T to the gravitational binding energy W exceeds a critical ratio $T/W > 0.27$ [35]. This instability will deform the star into a bar shape, which will emit gravitational waves in the kHz frequency band. The instability of a star with lower rate of rotation is driven by gravitation when the surface waves are counter-rotating in the rotating frame of the object but co-rotating with the star in the frame of a distant observer. In this case, positive angular momentum is removed from the star and the momentum of the mode becomes increasingly negative. This is known as the Chandrasekhar-Friedman-Schutz (CFS) instability. For f - and p - modes this instability is relevant at high frequencies close to the breakup-limit of the star, and, thus only partially interesting as a source of gravitational waves [11]. The r -modes, however, satisfy the CFS instability at arbitrary frequencies. The expected emission frequency for GWs from r -modes is $f \approx \frac{4}{3}\nu$ [36]. While it is expected that the amplitude of the waves is weaker than the amplitude of burst events, the time scale of emission is uncertain, ranging from few months after the birth of the neutron star to several thousands of years [36].

Free precession

A spinning neutron star in free precession, which is the rotation of the angular velocity vector around the axis of the angular momentum, is expected to produce gravitational

waves of amplitude [32]

$$h_0 \sim 10^{-27} \left(\frac{\theta_w}{0.1 \text{ rad}} \right) \left(\frac{1 \text{ kpc}}{d} \right) \left(\frac{\nu}{500 \text{ Hz}} \right)^2, \quad (1.35)$$

where θ_w is the “wobble” angle. The maximal wobble angle supported by the crust is expressed as [37]

$$\theta_w^{\text{max}} \simeq 1.8 \times 10^{-2} \left(\frac{500 \text{ Hz}}{\nu} \right)^2 \left(\frac{\sigma}{10^{-3}} \right), \quad (1.36)$$

where σ is the breaking strain of the crust. In [37], it has been shown that the gravitational wave frequency is expected at $\nu + \nu_{\text{prec}}$, 2ν and $2(\nu + \nu_{\text{prec}})$, where ν_{prec} is the precession frequency.

2. Searches for continuous gravitational waves

The main focus of the present work is the improvement of the searches for continuous gravitational waves from unknown sources performed in a distributed computing environment, e.g., Einstein@Home. Typical candidates for such unknown objects are neutron stars.

The theoretical prediction of neutron stars has been made by Baade and Zwicky in 1934 [38], while the discovery of the first radio pulsar (PSR B1919+21) by Jocelyn Bell Burnell had to wait until 1967 [39]. Neutron stars have mean density, which is two to three times the normal nuclear density $2.8 \times 10^{14} \text{ g cm}^{-3}$ in a radius $R \sim 10 \text{ km}$ with masses in the range one to two Solar masses, $M_{\odot} = 1.989 \times 10^{33} \text{ g}$, spinning with periods down to $P \sim 1.4 \text{ ms}$, and with large magnetic fields up to $B \lesssim 10^{15} \text{ G}$ [40].

In this chapter in Sec. 2.1 we discuss the signal model for a continuous wave emitter and in Sec. 2.2 a detection statistic, called $\mathcal{F}$ -statistic. In Sec. 2.3 we summarize some of the basic aspects of the $\mathcal{F}$ -statistic based searches.

2.1. Signal model

Consider the Earth, the Sun and a gravitational-wave source, as schematically presented in Fig. 2.1. The phase $\phi(t)$ of the expected signal from an isolated neutron star can be written as [41]:

$$\phi(t) = \phi_0 + 2\pi \sum_{k=0}^s \frac{f^{(k)}(\tau_{\text{ref}})}{(k+1)!} [\tau(t)]^{k+1} , \quad (2.1)$$

where ϕ_0 is the initial phase, $f^{(k)}$ is the k -th spindown, i.e., the k -th time derivative of the frequency at some fixed reference time τ_{ref} in the solar-system barycenter (SSB), and $\tau(t)$ is the arrival time of the wavefront in the SSB. Expressed as a function of the time t measured at the detector site, $\tau(t)$ yields ¹:

$$\tau(t) = t + \frac{\vec{r}(t) \cdot \vec{n}}{c} - \tau_{\text{ref}} , \quad (2.2)$$

¹In numerical computations, the right-hand-side of Eq. (2.15) is completed by the addition of the solar system Einstein delay term and subtraction of the Shapiro delay. The Einstein delay accounts for the effect of the gravitational redshift and time dilation due to the motions of the Earth and other bodies, while the Shapiro delay is due to the propagation of the wave in the curved space-time in the solar system [42].

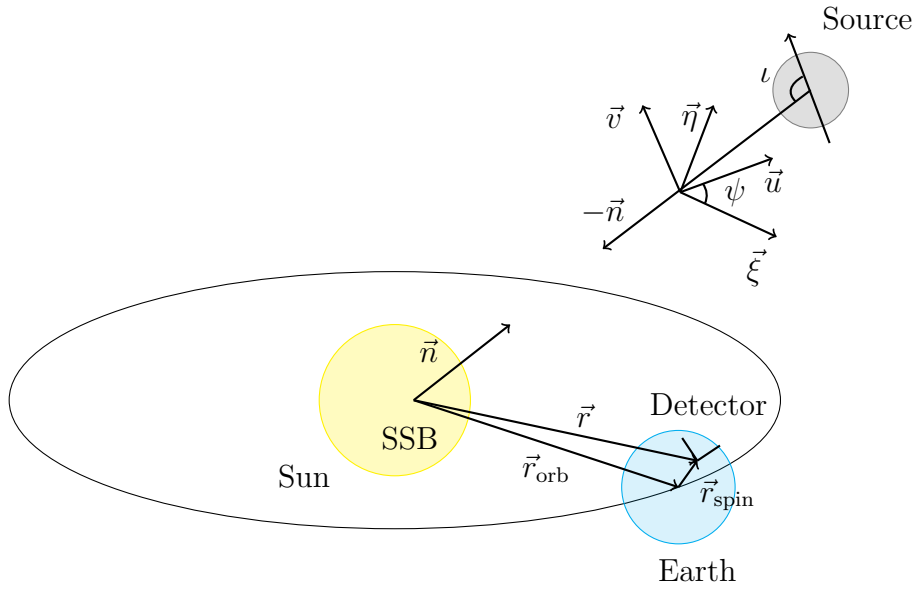


Figure 2.1.: The Sun, the Earth, a Detector, and a gravitational-wave Source: $\vec{r}$ points from the solar-system barycenter (SSB) to the detector, $\vec{r}_{orb}$ is the orbital vector pointing from the SSB to the center of the Earth, $\vec{r}_{spin}$ points from the Earth's center to the detector, $\vec{n}$ points to the source, ι is the inclination angle, and ψ is the polarization angle, i.e., the counter-clockwise measured rotation of the $\vec{u} - \vec{v}$ plane around $-\vec{n}$ with respect to the $\vec{\xi} - \vec{\eta}$ plane.

with the vector from the SSB to detector

$$\vec{r}(t) = \vec{r}_{\text{orb}}(t) + \vec{r}_{\text{spin}}(t) , \quad (2.3)$$

where the vector

$$\vec{n} = (\cos \delta \cos \alpha, \cos \delta \sin \alpha, \sin \delta) \quad (2.4)$$

expressed in equatorial coordinates, right ascension α and declination δ , points to the source. The orbital vector $\vec{r}_{\text{orb}}(t)$ points from the SSB to the center of the Earth, and $\vec{r}_{\text{spin}}$ points from Earth's center to the gravitational-wave detector. The relative motion of the detector causes frequency modulation of the received signal, which is well approximated by the Doppler relation [43]

$$\dot{f}(t) \approx f(\tau) \left(1 + \frac{\vec{v} \cdot \vec{n}}{c} \right) , \quad (2.5)$$

where $\dot{f}(t)$ is the instantaneous frequency at the detector, $f(\tau)$ is the SSB frequency of the wave, and $\vec{v}$ is the detector velocity. In addition, the signal undergoes amplitude modulation, caused by the antenna-patterns.

By convention a SSB fixed reference frame is constructed $\{\vec{\xi}, \vec{\eta}, -\vec{n}\}$ with $\vec{\xi} \equiv \vec{n} \times \vec{z} / |\vec{n} \times \vec{z}|$ and $\vec{\eta} \equiv \vec{\xi} \times \vec{n}$ [44]; see Fig. 2.1. Denoting with $\{\vec{u}, \vec{v}, -\vec{n}\}$ the orthonormal basis of the wave plane, the polarization angle ψ describes the counter-clockwise measured rotation of the $\vec{u} - \vec{v}$ plane around $-\vec{n}$ with respect to the $\vec{\xi} - \vec{\eta}$ plane, i.e.,

$$\begin{pmatrix} \vec{u} \\ \vec{v} \end{pmatrix} = \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix} \begin{pmatrix} \vec{\xi} \\ \vec{\eta} \end{pmatrix} . \quad (2.6)$$

With $\vec{z} = (0, 0, 1)$, $\vec{\xi}$ and $\vec{\eta}$ are explicitly given by

$$\vec{\xi} = (\sin \alpha, -\cos \alpha, 0) \quad \text{and} \quad \vec{\eta} = (-\cos \alpha \sin \delta, -\sin \alpha \sin \delta, \cos \delta) . \quad (2.7)$$

We define the polarization-independent basis tensors in the wave-frame [45]:

$$\epsilon_{+ \ ij} \equiv \xi_i \xi_j - \eta_i \eta_j \quad \text{and} \quad \epsilon_{\times \ ij} \equiv \xi_i \eta_j + \eta_i \xi_j , \quad (2.8)$$

which are explicitly written as

$$\epsilon_{+}^{\vec{r}} = \begin{pmatrix} \sin^2 \alpha - \cos^2 \alpha \sin^2 \delta & -\frac{1}{2} \sin 2\alpha (1 + \sin^2 \delta) & \frac{1}{2} \cos \alpha \sin 2\delta \\ -\frac{1}{2} \sin 2\alpha (1 + \sin^2 \delta) & \cos^2 \alpha - \sin^2 \alpha \sin^2 \delta & \frac{1}{2} \sin \alpha \sin 2\delta \\ \frac{1}{2} \cos \alpha \sin 2\delta & \frac{1}{2} \sin \alpha \sin 2\delta & -\cos^2 \delta \end{pmatrix} , \quad (2.9)$$

and

$$\epsilon_{\times}^{\vec{r}} = \begin{pmatrix} -\sin 2\alpha \sin \delta & (\cos^2 \alpha - \sin^2 \alpha) \sin \delta & \sin \alpha \cos \delta \\ (\cos^2 \alpha \sin^2 \alpha) \sin \delta & \sin 2\alpha \sin \delta & -\cos \alpha \cos \delta \\ \sin \alpha \cos \delta & -\cos \alpha \cos \delta & 0 \end{pmatrix} , \quad (2.10)$$

The detector tensor for an interferometer with arms along $\vec{l}_1$ and $\vec{l}_2$ is defined as [44]

$$d_{ij} = \frac{1}{2}(l_{1i}l_{1j} - l_{2i}l_{2j}) . \quad (2.11)$$

With Eqs. (2.9), (2.10) and (2.11) we define the amplitude modulation functions $a^X(t)$ and $b^X(t)$ at detector X :

$$a^X(t, \vec{n}) \equiv d^X(t)_{ij}\epsilon_+^{ij} \quad \text{and} \quad b^X(t, \vec{n}) \equiv d^X(t)_{ij}\epsilon_\times^{ij} , \quad (2.12)$$

which are related to the antenna pattern functions F_+ and $F_\times$ [Eq. (1.21)] through

$$F_+(t) = a^X(t) \cos 2\psi + b^X(t) \sin 2\psi , \quad (2.13)$$

$$F_\times(t) = -a^X(t) \sin 2\psi + b^X(t) \cos 2\psi . \quad (2.14)$$

For example, for the particular orientation of the detector arms $\vec{l}_1 = (1, 0, 0)$ and $\vec{l}_2 = (0, 1, 0)$, Eq. (2.11) yields

$$d = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 \end{pmatrix} ; \quad (2.15)$$

thus the Eqs. in (2.12) take the form

$$a = \frac{1}{2}(\sin^2 \alpha - \cos^2 \alpha)(1 + \sin^2 \delta) \quad \text{and} \quad b = -\sin 2\alpha \sin \delta . \quad (2.16)$$

As shown in [28, 46] we can express the dimensionless strain signal $h^X(t)$ of a continuous gravitational wave at detector X as a sum of four detector independent signal-amplitudes $\mathcal{A}^\mu$ and four detector dependent basis waveforms $h_\mu^X(t)$:

$$h^X(t) = \sum_{\mu=1}^4 \mathcal{A}^\mu h_\mu^X(t) , \quad (2.17)$$

where

$$\mathcal{A}^1 = A_+ \cos \phi_0 \cos 2\psi - A_\times \sin \phi_0 \sin 2\psi , \quad (2.18)$$

$$\mathcal{A}^2 = A_+ \cos \phi_0 \sin 2\psi + A_\times \sin \phi_0 \cos 2\psi , \quad (2.19)$$

$$\mathcal{A}^3 = -A_+ \sin \phi_0 \cos 2\psi - A_\times \cos \phi_0 \sin 2\psi , \quad (2.20)$$

$$\mathcal{A}^4 = -A_+ \sin \phi_0 \sin 2\psi + A_\times \cos \phi_0 \cos 2\psi , \quad (2.21)$$

and

$$h_1^X(t) = a^X(t) \cos \phi^X(t) , \quad (2.22)$$

$$h_2^X(t) = b^X(t) \cos \phi^X(t) , \quad (2.23)$$

$$h_3^X(t) = a^X(t) \sin \phi^X(t) , \quad (2.24)$$

$$h_4^X(t) = b^X(t) \sin \phi^X(t). \quad (2.25)$$

For a triaxial neutron star, the two polarization amplitudes A_+ and $A_\times$ written in terms of the overall amplitude h_0 and inclination angle ι are:

$$A_+ = \frac{1}{2} h_0 (1 + \cos^2 \iota), \quad (2.26)$$

$$A_\times = h_0 \cos \iota. \quad (2.27)$$

We refer to the set of signal parameters $\lambda_s \equiv \{\alpha, \delta, f^{(k)}\}$ as *phase-evolution* parameters, while the set $\{h_0, \iota, \psi, \phi_0\}$ is known as *amplitude* parameters. Depending on what fraction of the signal parameters λ_s are known, we distinguish three types of searches, namely: *targeted*, when λ_s is known; *directed*, when only the sky position is known; and *wide* parameter-space searches, when λ_s is completely unknown. Targeted searches are typically performed with the heterodyne method [47], where Bayesian analysis is used to infer the unknown amplitude parameters. For directed and wide parameter-space searches a maximum-likelihood detection statistic, $\mathcal{F}$ -statistic [28, 46] is usually applied. These are the type of searches that we mainly consider in this work.

2.2. Detection statistic

In the previous section we introduced the signal model. In the following, we discuss how to detect a signal $h(t; \mathcal{A}, \lambda_s)$, if it is present in the data collected from a gravitational-wave detector. The classical detection problem involves the computation of a likelihood ratio [4], and its comparison to a threshold in order to confirm or discard a hypothesis.

2.2.1. Hypothesis testing

Let the detector data be $x(t)$ and a signal $h(t; \mathcal{A}, \lambda_s)$ be described by Eq. (2.17). We aim to test the two hypothesis; the data is noise:

$$\mathcal{H}_n : x(t) = n(t), \quad (2.28)$$

or there is a signal in the data

$$\mathcal{H}_s : x(t) = n(t) + h(t; \mathcal{A}, \lambda_s). \quad (2.29)$$

According to the Neyman-Pearson lemma, the optimal test is given by the likelihood ratio

$$\Lambda(h) \equiv \frac{\text{pdf}(x|\mathcal{A}, \lambda_s)}{\text{pdf}(x|0)}, \quad (2.30)$$

where $\text{pdf}(x|\mathcal{A}, \lambda_s)$ is the probability density function to observe the data when the signal is present and $\text{pdf}(x|0)$ is the probability density function when the signal is absent. To decide about the signal hypothesis $\mathcal{H}_s$, we compare the value of the likelihood ratio Λ

to a threshold Λ_{th} . We decide to accept $\mathcal{H}_s$, if the threshold is crossed, i.e., $\Lambda(h) > \Lambda_{\text{th}}$. Setting the threshold Λ_{th} and making a decision is subject to two possible errors; namely, false-alarm and false-dismissal.

The false-alarm probability, denoted with p_{fA} , is the probability to falsely decide that hypothesis $\mathcal{H}_s$ is true, while the true hypothesis is instead $\mathcal{H}_n$, i.e.,

$$p_{\text{fA}}(\Lambda_{\text{th}}) \equiv P(\Lambda(h) > \Lambda_{\text{th}} | \mathcal{H}_n) = \int_{\Lambda_{\text{th}}}^{\infty} d\Lambda \text{pdf}(\Lambda | \mathcal{H}_n) . \quad (2.31)$$

Therefore if we fix the false-alarm probability to a certain value, p_{fA}^* , and we know the probability density function of the noise hypothesis $\text{pdf}(\Lambda(h) | \mathcal{H}_n)$, we can integrate the right-hand-side of Eq. (2.31) and obtain a threshold $\Lambda_{\text{th}}(p_{\text{fA}}^*)$.

The false-dismissal probability, denoted p_{fD} , is the probability to decide that hypothesis $\mathcal{H}_n$ is true, while the true hypothesis is $\mathcal{H}_s$, i.e.,

$$p_{\text{fD}}(\Lambda_{\text{th}}, h) \equiv P(\Lambda(h) < \Lambda_{\text{th}} | \mathcal{H}_s) = \int_{-\infty}^{\Lambda_{\text{th}}} d\Lambda \text{pdf}(\Lambda | \mathcal{H}_s) . \quad (2.32)$$

Using the false-dismissal probability we can easily compute the detection probability, i.e.,

$$p_{\text{det}}(\Lambda_{\text{th}}, h) \equiv P(\Lambda(h) > \Lambda_{\text{th}} | \mathcal{H}_s) = \int_{\Lambda_{\text{th}}}^{\infty} \text{pdf}(\Lambda | \mathcal{H}_s) , \quad (2.33)$$

therefore

$$p_{\text{det}}(\Lambda_{\text{th}}, h) = 1 - p_{\text{fD}}(\Lambda_{\text{th}}, h) . \quad (2.34)$$

2.2.2. The $\mathcal{F}$ -statistic

The $\mathcal{F}$ -statistic, first derived in [28] and extended to a multi-detector statistic in [46], is the log-likelihood ratio $\ln \Lambda(x, \mathcal{A}, \lambda_s)$ maximized over the amplitude parameters $\mathcal{A}$, and therefore only a function of the phase-evolution parameters λ_s , namely

$$\mathcal{F}(x, \lambda_s) \equiv \max_{\mathcal{A}} (\ln \Lambda(x, \mathcal{A}, \lambda_s)) \quad (2.35)$$

To discuss some important aspects of the $\mathcal{F}$ -statistic based searches in the next section, it will be useful to evaluate the right-hand-side of Eq. (2.35).

Following [48], we begin the evaluation by noticing, that the conditional probability to observe the data given that the signal is present equals the probability to observe the data without the signal, given that the signal is not present, i.e.,

$$\text{pdf}(x | \mathcal{A}, \lambda_s) = \text{pdf}(x - h | 0) . \quad (2.36)$$

With the above equation, the likelihood ratio yields

$$\Lambda(h) \equiv \frac{\text{pdf}(x - h | 0)}{\text{pdf}(x | 0)} . \quad (2.37)$$

Next recall that $n(t)$ denotes a stationary random process with zero mean; see Eqs. (1.23) and (1.24). Assuming that each data sample x_i follows the distribution of $n(t)$, i.e., Gaussian,

$$\text{pdf}(x_i|0) \propto \exp\left(-\frac{1}{2}C^{-1}x_i^2\right), \quad (2.38)$$

thus the joint probability is a product of Gaussians, which leads to a sum in the exponential function, namely

$$\text{pdf}(x|0) \propto \exp\left(-\frac{1}{2}\sum_{j,k=1}^S C_{jk}^{-1}x_jx_k\right), \quad (2.39)$$

where C_{jk}^{-1} is the inverse of the correlation function [49]

$$\begin{aligned} C(\tau) &= \langle x(t)x(t+\tau) \rangle \\ &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T dt x(t)x(t+\tau) \end{aligned} \quad (2.40)$$

defined such that

$$\delta_{jk} \equiv \sum_{l=1}^S C((j-l)\Delta t) C^{-1}((l-k)\Delta t), \quad (2.41)$$

and S is the number of samples. For convenience we map the discretely taken data to the continuum through

$$x_i = x(t_i), \quad t_i - t_j = (i-j)\Delta t, \quad \text{and} \quad \Delta t = \frac{T}{S-1}. \quad (2.42)$$

Using

$$\delta(t_j - t_k) = \lim_{\substack{\Delta t \rightarrow 0 \\ T \rightarrow \infty}} \frac{1}{\Delta t} \delta_{jk} \quad (2.43)$$

it is useful to show the following relation [48]:

$$\begin{aligned} e^{2\pi i f t_k} &= \sum_j e^{2\pi i f t_j} \delta_{ij} \\ &= \lim_{\substack{\Delta t \rightarrow 0 \\ T \rightarrow \infty}} \frac{1}{\Delta t^2} \sum_j \Delta t e^{2\pi i f t_j} \sum_l \Delta t C(t_j - t_l) C^{-1}(t_l - t_k) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t^2} \int_{-\infty}^{\infty} dt_j e^{2\pi i f t_j} \int_{-\infty}^{\infty} dt_l C(t_j - t_l) C^{-1}(t_l - t_k) \\ &\quad | \quad \text{with} \quad \tau = t_j - t_l \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t^2} \int_{-\infty}^{\infty} dt_l e^{2\pi i f t_l} C^{-1}(t_l - t_k) \int_{-\infty}^{\infty} d\tau e^{2\pi i f \tau} C(\tau) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t^2} \frac{1}{2} S_n(f) \int_{-\infty}^{\infty} dt_l e^{2\pi i f t_l} C^{-1}(t_l - t_k), \end{aligned} \quad (2.44)$$

where we applied the Wiener-Khintchine relation [3]

$$\frac{1}{2}S_n(f) \equiv \int_{-\infty}^{\infty} d\tau e^{2\pi i f \tau} C(\tau) \quad (2.45)$$

to express the correlation function through the PSD $S_n(f)$. Defining

$$\widetilde{C}^{-1}(f, t_k) \equiv \int_{-\infty}^{\infty} dt_l C^{-1}(t_l - t_k) e^{2\pi i f t_l} \quad (2.46)$$

from Eq. (2.44) we finally obtain

$$\widetilde{C}^{-1}(f, t_k) = \lim_{\Delta t \rightarrow 0} \Delta t^2 \frac{2e^{2\pi i f t_k}}{S_n(f)} . \quad (2.47)$$

With the above equation and using Parseval's theorem [Eq.(1.24)] we can evaluate the continuum limit of the sum in Eq. (2.39)[48]:

$$\begin{aligned} \lim_{\substack{\Delta t \rightarrow 0 \\ T \rightarrow \infty}} \sum_{j,k=1}^S C_{jk}^{-1} x_j x_k &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t^2} \sum_{j,k=1}^S \Delta t^2 C^{-1}(t_j - t_k) x(t_j) x(t_k) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dt_j dt_k C^{-1}(t_j - t_k) x(t_j) x(t_k) \\ &= \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t^2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} df dt_k \widetilde{C}^{-1}(f, t_k) \widetilde{x}^*(f) x(t_k) \\ &= 2 \int_{-\infty}^{\infty} df \frac{\widetilde{x}^*(f)}{S_n(|f|)} \int_{-\infty}^{\infty} dt_k e^{2\pi i f t_k} x(t_k) \\ &= 2 \int_{-\infty}^{\infty} df \frac{\widetilde{x}(f) \widetilde{x}^*(f)}{S_n(|f|)} . \end{aligned} \quad (2.48)$$

With the right-hand-side of Eq. (2.48) we define the scalar product for two real functions x and y , namely

$$\langle x|y \rangle \equiv 4 \int_0^{\infty} df \frac{\widetilde{x}(f) \widetilde{y}^*(f)}{S_n(f)} . \quad (2.49)$$

The scalar product from the above equation is the matched filter of the two functions x and y . Using the matched filter, Eq. (2.39) yields

$$\text{pdf}(x|0) \propto \exp \left(-\frac{1}{2} \langle x|x \rangle \right) , \quad (2.50)$$

and we obtain the likelihood ratio from Eq. (2.30)

$$\begin{aligned} \Lambda(h) &= \frac{P(x-h|0)}{P(x|0)} \\ &= \exp \left(\langle x|h \rangle - \frac{1}{2} \langle h|h \rangle \right) . \end{aligned} \quad (2.51)$$

Taking the logarithm of the above equation and using Eq. (2.17), where we omit the sum and use the summation convention, we obtain

$$\begin{aligned}
\ln \Lambda &= \langle x|h \rangle - \frac{1}{2} \langle h|h \rangle \\
&= \langle x|\mathcal{A}^\mu h_\mu \rangle - \frac{1}{2} \langle \mathcal{A}^\mu h_\mu | \mathcal{A}^\nu h_\nu \rangle \\
&= A^\mu \langle x|h_\mu \rangle - \frac{1}{2} \mathcal{A}^\mu \mathcal{A}^\nu \langle h_\mu | h_\nu \rangle \\
&= \mathcal{A}^\mu x_\mu - \frac{1}{2} \mathcal{A}^\mu \mathcal{A}^\nu \mathcal{M}_{\mu\nu} ,
\end{aligned} \tag{2.52}$$

with

$$x_\mu \equiv \langle x|h_\mu \rangle \quad \text{and} \quad \mathcal{M}_{\mu\nu} \equiv \langle h_\mu | h_\nu \rangle . \tag{2.53}$$

Maximization of Eq. (2.52) with respect to the unknown amplitudes yields the four maximum-likelihood amplitudes

$$\mathcal{A}_{\text{ML}}^\mu = \mathcal{M}^{\mu\nu} x_\nu . \tag{2.54}$$

By substitution of the maximum-likelihood amplitudes in Eq. (2.48) we can write the $\mathcal{F}$ -statistic in the form

$$2\mathcal{F}(x, \lambda_s) = x_\mu \mathcal{M}^{\mu\nu} x_\nu . \tag{2.55}$$

For a network of detectors, assuming that the noise of the individual detectors is independent from each other, the multidetector scalar product is the sum over detectors X of Eq. (2.49), i.e., the sum over the single detector scalar product, namely

$$\langle x|y \rangle = \sum_X \langle x^X | y^X \rangle . \tag{2.56}$$

2.3. $\mathcal{F}$ -statistic based searches

The $\mathcal{F}$ -statistic, $2\mathcal{F}$, follows a non-central χ^2 distribution with four degrees of freedom and a non-centrality parameter that equals the squared signal-to-noise (SNR) ratio

$$\rho^2 \equiv \langle h|h \rangle . \tag{2.57}$$

The expectation value of the $\mathcal{F}$ -statistic is therefore

$$E[2\mathcal{F}] = 4 + \rho^2 , \tag{2.58}$$

with variance

$$\sigma^2[2\mathcal{F}] = 2(4 + 2\rho^2) . \tag{2.59}$$

Dividing the observed data in N number of segments with duration ΔT , we can compute a new statistic Σ , referred to as semicoherent $\mathcal{F}$ -statistic (StackSlide). Denoting the index over segments $k = 1 \dots N$, the StackSlide statistic Σ is defined as

$$\Sigma \equiv \sum_{k=1}^N 2\mathcal{F}_k, \quad (2.60)$$

i.e., a simple sum of $\mathcal{F}$ -statistic values $2\mathcal{F}_k$ computed across all N segments. In the presence of signal, the expected value is

$$E[\Sigma] = 4N + \rho_\Sigma^2 \quad (2.61)$$

with variance

$$\sigma^2[\Sigma] = 2(4N + 2\rho_\Sigma^2), \quad (2.62)$$

where ρ_Σ^2 is

$$\rho_\Sigma^2 \equiv \sum_{k=1}^N \rho_k^2. \quad (2.63)$$

To illustrate Eqs. (2.31) and (2.32) in Fig. 2.2 we plot the central χ^2 distribution with four degrees of freedom and the non-central χ^2 distribution with four degrees of freedom and non-centrality parameter $\rho^2 = 16$.

2.3.1. Signal-to-noise ratio

In this subsection we evaluate Eq. (2.57) to obtain a useful formula relating the SNR^2 to the intrinsic gravitational-wave amplitude h_0 , and to the amount, and the noise level of the data. Before proceeding, we notice the following useful facts. First, assuming a quasi monochromatic signal (in the SSB), the main contribution to the scalar product, Eq. (2.49), comes from a narrow band around the signal frequency, thus S_n is approximately constant. This allows us to replace the integration over frequency in Eq. (2.49) with a time integral, namely [11]:

$$\langle x|y \rangle \approx \frac{2}{S_n(f_0)} \int_0^T dt x(t)y(t). \quad (2.64)$$

Second, the phase functions $\sin \phi(t)$ and $\cos \phi(t)$ in Eqs. (2.22) to (2.25) are periodic on a time scale very short compared to the observation time, and averaging over one period yields

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} d\phi \sin^2(\phi) &= \frac{1}{2}, \\ \frac{1}{2\pi} \int_0^{2\pi} d\phi \cos^2(\phi) &= \frac{1}{2}, \\ \frac{1}{2\pi} \int_0^{2\pi} d\phi \sin(\phi) \cos(\phi) &= 0. \end{aligned} \quad (2.65)$$

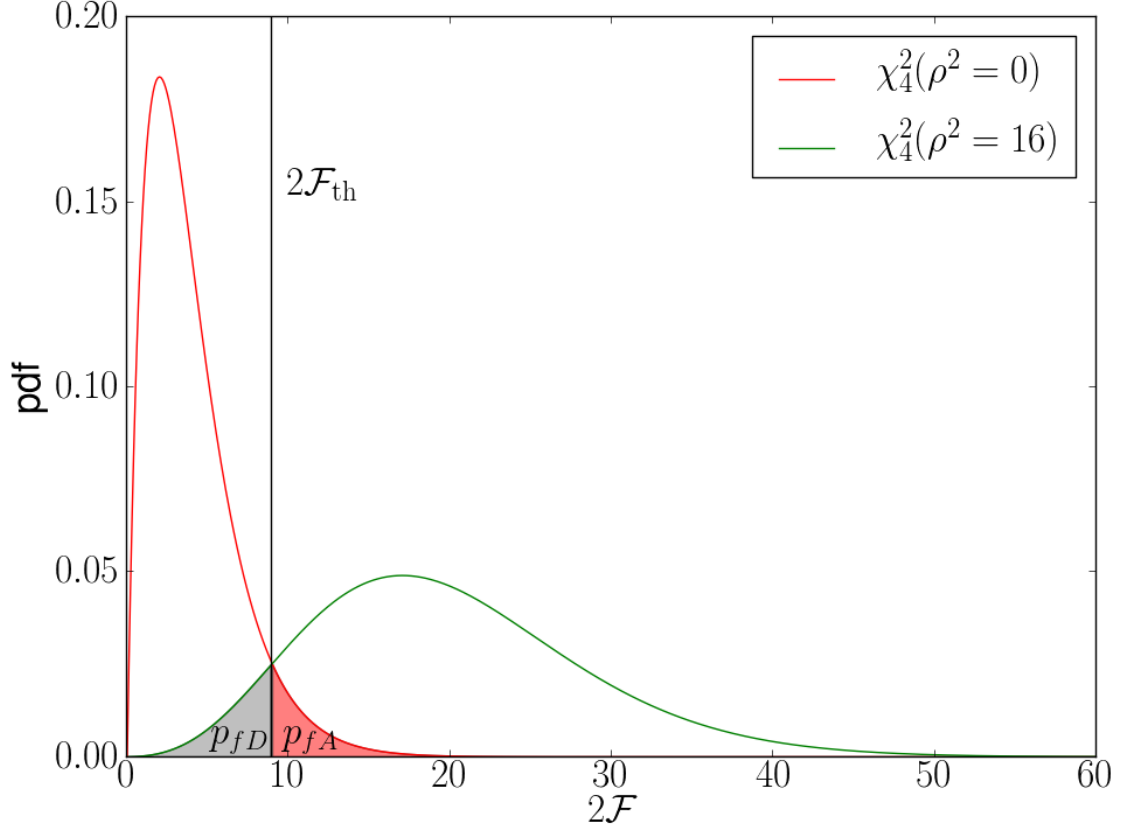


Figure 2.2.: Probability density for a χ^2 distribution with four degrees of freedom and different non-centrality parameters. The red curve ($\chi_4^2(\rho^2 = 0)$) is the pdf of the noise hypothesis $\mathcal{H}_n$, and the green curve ($\chi_4^2(\rho^2 = 16)$) is the pdf of the signal hypothesis $\mathcal{H}_s$. In this example the intersection of the two pdfs is taken to be the threshold $2\mathcal{F}_{\text{th}}$. The integral of the red shaded area is the false-alarm probability, Eq. (2.31), and the integral of the gray shaded area is the false-dismissal probability, Eq. (2.32).

We continue now with the computation of the SNR², which can be written as

$$\rho^2 = \mathcal{A}^\mu \mathcal{M}_{\mu\nu} \mathcal{A}^\nu, \quad (2.66)$$

with

$$\mathcal{M}_{\mu\nu} = \sum_X \langle h_\mu^X(t) | h_\nu^X(t) \rangle \approx \frac{1}{2} \begin{pmatrix} \mathcal{M}_a & \mathcal{M}_c & 0 & 0 \\ \mathcal{M}_c & \mathcal{M}_b & 0 & 0 \\ 0 & 0 & \mathcal{M}_a & \mathcal{M}_c \\ 0 & 0 & \mathcal{M}_c & \mathcal{M}_b \end{pmatrix}, \quad (2.67)$$

where the approximation is due to the usage of Eqs. 2.65. With this, the squared signal to noise ratio becomes

$$\rho^2 = \frac{1}{2} ((\mathcal{A}_1^2 + \mathcal{A}_3^2) \mathcal{M}_a + (\mathcal{A}_2^2 + \mathcal{A}_4^2) \mathcal{M}_b + 2(\mathcal{A}_1 \mathcal{A}_2 + \mathcal{A}_3 \mathcal{A}_4) \mathcal{M}_c). \quad (2.68)$$

Further calculation yields

$$\mathcal{A}_1^2 + \mathcal{A}_3^2 = \frac{h_0^2}{4} ((1 + \cos^2 \iota)^2 \cos^2 2\psi + 4 \cos^2 \iota \sin^2 2\psi) = h_0^2 \alpha_1, \quad (2.69)$$

$$\mathcal{A}_2^2 + \mathcal{A}_4^2 = \frac{h_0^2}{4} ((1 + \cos^2 \iota)^2 \sin^2 2\psi + 4 \cos^2 \iota \cos^2 2\psi) = h_0^2 \alpha_2, \quad (2.70)$$

$$\mathcal{A}_1 \mathcal{A}_2 + \mathcal{A}_3 \mathcal{A}_4 = \frac{h_0^2}{4} (1 - \cos^2 \iota)^2 \sin 2\psi \cos 2\psi = h_0^2 \alpha_3, \quad (2.71)$$

and thus

$$\rho^2 = \frac{h_0^2}{2} (\alpha_1 \mathcal{M}_a + \alpha_2 \mathcal{M}_b + 2\alpha_3 \mathcal{M}_c). \quad (2.72)$$

At this point we take into account, that the data is not continuous but discretized in time intervals Δt . Further, the data is usually preprocessed and available in the form of Short-baseline Fourier Transforms (SFTs). The baseline of the SFTs is chosen, such that a putative signal will appear monochromatic during the SFT duration [50]. The length of the SFTs used in this work is $T_{\text{SFT}} = 1800$ s. Using the SFT duration and time discretization Δt we can express a point in time as

$$t_{n,m} = (n-1)T_{\text{SFT}} + m\Delta t, \quad (2.73)$$

where $n \in [0, N_{\text{SFT}}]$ is an index pointing the number of the SFT, N_{SFT} is the total number of SFTs, and $m \in [0, M]$ is the number of steps in a SFT, thus $T_{\text{SFT}} = M\Delta t$. We denote the number of SFTs per detector with N_{SFT}^X .

We first compute

$$\begin{aligned}
\mathcal{M}_a &= 2 \sum_X \langle h_1^X(t) | h_1^X(t) \rangle \\
&= 4 \sum_X \int_{t_0}^{t_1} \frac{h_1^X(t) h_1^X(t)}{S^X} \\
&= 4 \sum_X \sum_{l=1}^{MN_{\text{SFT}}^X} \frac{h_{1l}^X h_{1l}^X}{S_l^X} \Delta t \\
&= 4 \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{a_n^X a_n^X}{S_n^X} \sum_{m=1}^M \cos^2 \phi_{nj}^X \Delta t \\
&\approx 2 \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{a_n^X a_n^X}{S_n^X} M \Delta t \\
&= 2T_{\text{SFT}} \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{(a_n^X)^2}{S_n^X}
\end{aligned} \tag{2.74}$$

In a similar way we obtain

$$\mathcal{M}_b = 2T_{\text{SFT}} \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{(b_n^X)^2}{S_n^X} \tag{2.75}$$

and

$$\mathcal{M}_c = 2T_{\text{SFT}} \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{a_n^X b_n^X}{S_n^X} . \tag{2.76}$$

Averaging over $\cos \iota$ and ψ yields:

$$\langle \alpha_1 \rangle_{\cos \iota, \psi} = \frac{1}{2} \int_{-1}^1 d \cos \iota \frac{1}{2\pi} \int_{-\pi/4}^{\pi/4} \alpha_1 = \frac{2}{5} , \tag{2.77}$$

$$\langle \alpha_2 \rangle_{\cos \iota, \psi} = \frac{1}{2} \int_{-1}^1 d \cos \iota \frac{1}{2\pi} \int_{-\pi/4}^{\pi/4} \alpha_2 = \frac{2}{5} , \tag{2.78}$$

$$\langle \alpha_3 \rangle_{\cos \iota, \psi} = \frac{1}{2} \int_{-1}^1 d \cos \iota \frac{1}{2\pi} \int_{-\pi/4}^{\pi/4} \alpha_3 = 0 . \tag{2.79}$$

Putting everything together the SNR^2 averaged over polarization and inclination angle

becomes

$$\begin{aligned}
\langle \rho^2 \rangle_{\cos \iota, \psi} &= \frac{2h_0^2}{5} \left(T_{\text{SFT}} \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{(a_n^X)^2}{S_n^X} + T_{\text{SFT}} \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{(b_n^X)^2}{S_n^X} \right) \\
&= \frac{2h_0^2 T_{\text{SFT}}}{5} \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{(a_n^X)^2 + (b_n^X)^2}{S_n^X} \\
&= \frac{2h_0^2}{5} \sum_X \sum_{n=1}^{N_{\text{SFT}}^X} \frac{(a_n^X)^2 + (b_n^X)^2}{S_n^X} \frac{T_{\text{data}}^X}{N_{\text{SFT}}^X}
\end{aligned} \tag{2.80}$$

where we define the amount of data per detector X to be

$$T_{\text{data}}^X = N_{\text{SFT}}^X T_{\text{SFT}} . \tag{2.81}$$

For a directed search, further evaluation of Eq. (2.80) should be done numerically. However it is useful to have a more general form, independent of the sky position of the source. Thus assuming the particular form of the antenna-pattern functions given in Eq. (2.16) and averaging isotropically over the sky, yields

$$\langle a^2 \rangle_{\vec{n}} = \frac{7}{30}, \quad \text{and} \quad \langle b^2 \rangle_{\vec{n}} = \frac{1}{6}. \tag{2.82}$$

So we finally obtain the SNR^2 averaged over sky position, $\cos \iota$ and ψ , namely [44]

$$\langle \rho^2 \rangle_{\vec{n}, \cos \iota, \psi} = \frac{4}{25} h_0^2 T_{\text{data}} \mathcal{S}^{-1}(f_0), \tag{2.83}$$

where T_{data} is all of the data, i.e., $T_{\text{data}} \equiv N_{\text{SFT}} T_{\text{SFT}}$ and

$$\mathcal{S}^{-1}(f_0) \equiv N_{\text{SFT}}^{-1} \sum_{n=1}^{N_{\text{SFT}}} S_n^{-1}(f_0). \tag{2.84}$$

is the harmonic mean of PSDs over all SFTs.

2.3.2. Wide parameter-space searches

Up to now we implicitly assumed that the phase-evolution parameters of the continuous gravitational-wave signal are known. However searching for emission from previously unknown objects implies that at least a part of the phase-evolution parameters are unknown. This introduces a layer of complexity with a very large impact on the computing cost and sensitivity of the search. Without knowledge of the signal parameters, the search consists in the computation of the detection statistic, the matched filter, in “every” point of the parameter space. The question is therefore, how can we cover the parameter space efficiently, or in other words, what does “every” means?

To give a quantitative answer to the question stated above, we use the concept of mismatch, which was introduced in [51, 52]. By definition, the mismatch between an arbitrary template λ and the true signal location λ_s is the fractional loss of SNR², namely

$$\mu \equiv \frac{\rho^2(\lambda_s) - \rho^2(\lambda)}{\rho^2(\lambda_s)}. \quad (2.85)$$

The mismatch can be expanded for a small offset $\Delta\lambda = \lambda - \lambda_s$ between the arbitrary parameter space point and the signal² [53]

$$\begin{aligned} \mu(\Delta\lambda) &= \mu(0) + \partial_i \mu(\Delta\lambda) \Delta\lambda^i \\ &+ \frac{1}{2} \partial_i \partial_j \mu(\Delta\lambda) \Delta\lambda^i \Delta\lambda^j + \mathcal{O}(\Delta\lambda^3), \end{aligned} \quad (2.86)$$

where $\partial_i \equiv \frac{\partial}{\partial \lambda^i}$. Noticing that the first two terms of the expansion vanish as μ is minimized at $\Delta\lambda = 0$, for small offsets we can write

$$\mu = g_{ij}(\lambda_s) \Delta\lambda^i \Delta\lambda^j + \mathcal{O}(\Delta\lambda^3). \quad (2.87)$$

where g_{ij} denotes the parameter-space metric. The metric g_{ij} is the quadratic description of the parameter-space in the vicinity of the signal, and can be used to measure distance, i.e.

$$\mu^*(\Delta\lambda) = g_{ij}(\lambda_s) \Delta\lambda^i \Delta\lambda^j. \quad (2.88)$$

With this, “every” means that the search points of the parameter space, the templates, are placed densely enough, so that the mismatch of a putative signal to the closest template is m at worst. The construction of an optimal template bank, which contain the smallest possible number of templates, has been studied in detail in [54]. It has been shown, that the A_n^* lattice-covering is the best for $n \leq 5$ and close to the best for dimensions $n \leq 16$, while the hyper-cubic lattice covering is still acceptable at lower dimensions $n \leq 4$.

The required number of templates $\mathcal{N}$ to cover the parameter space $\mathbb{P}$ is given by the general expression [54, 55]

$$\mathcal{N} = \theta_n m^{-n/2} \mathcal{V}_n, \quad \text{with} \quad \mathcal{V}_n \equiv \int_{\mathbb{T}_n} d^n \lambda \sqrt{\det g}, \quad (2.89)$$

where θ_n is the normalized thickness, m is the maximal-mismatch parameter, $\det g$ is the determinant of the corresponding parameter-space metric g_{ij} , and $\mathcal{V}_n$ denotes the metric volume of the n -dimensional space $\mathbb{T}_n \subseteq \mathbb{P}$ spanned by the template-bank. The normalized thickness corresponds to the number of templates per unit volume [54]. To give a rough estimate of the number of templates and therefore of the computing cost scalings of a $\mathcal{F}$ -statistic search, we will use Eq. (2.89), in particular $\sqrt{\det g}$, therefore we need to compute the metric g .

²We use summation convention with respect to repeated indices.

The metric of the multidetector $\mathcal{F}$ -statistic has been studied in [41] in great detail, and a simplified metric, called “phase metric” was introduced in [53]. Analytic expressions for a semicoherent metric can be found in [56]. Given a phase model, the metric components can be obtained as [53]

$$g_{ij}(\lambda) \equiv \langle \partial_i \phi \partial_j \phi \rangle - \langle \partial_i \phi \rangle \langle \partial_j \phi \rangle , \quad (2.90)$$

where $\langle . \rangle$ denotes time averaging over the observation time. Using the phase model Eq. (2.1) and including up to first-order spindown (4D parameter space) the dominant scalings with the observation time are [11]

$$g_{\theta\theta} \propto f^2 T^2 (v/c)^2, \quad g_{ff} \propto T^2, \quad g_{f\dot{f}} \propto T^4 , \quad (2.91)$$

where θ is the angular separation on the sky and $v/c \sim 10^{-4}$ is the maximal Doppler shift due to the orbital velocity v . With the above diagonal elements of g , neglecting possible off-diagonal elements, we obtain $\sqrt{\det g} \propto T^5 f^2$. With this and using Eq. (2.89), the number of templates scales as

$$\mathcal{N} \sim T^5 f^2 . \quad (2.92)$$

The computing cost per template is proportional to the observation time, and assuming data without gaps, we obtain the scaling of the computing cost, namely

$$C \sim T^6 f^2 . \quad (2.93)$$

This order of magnitude computing cost estimate shows that a year long fully coherent search over a wide parameter space, e.g., an all-sky survey in a kHz frequency range is computationally impossible [53].

The high computational requirement of the wide parameter-space searches requires the usage of distributed computing environments like Einstein@Home [57], where volunteers donate computing power by performing a search over a small region of the total parameter space. Searches done in the past include the Einstein@Home search for periodic gravitational waves in LIGO S4 [58] and LIGO S5 [59] data.

2.3.3. The choice of spin-down search band

As discussed in Chap. 4 the dimensionality of a wide parameter-space search, in particular the highest spin-down order that needs to be included in the search in order not to miss a signal, depends, i.a., on the size of the parameter-space in the given direction. We refer to this size as the search band. In this subsection, we discuss a possible way for an educated guess about the size of first- and higher-order spin-down search band based on the minimal age of the targeted sources and the braking indices.

The loss of rotational frequency of a pulsar can be written as a general power law [60]

$$\dot{\Omega} = -\mathcal{K}(n_1) \Omega^{n_1} , \quad (2.94)$$

where $\mathcal{K}(n_1)$ is a constant, and n_1 denotes the braking index. The braking index takes the value $n_1 = 3$ for purely dipolar magnetic-field emission, and $n_1 = 5$ for pure gravitational-wave emission [60]. Integration of Eq. (2.94) yields the age of the star

$$\tau = \frac{\Omega^{1-n_1} - \Omega_0^{1-n_1}}{(n_1 - 1)\mathcal{K}(n_1)} , \quad (2.95)$$

where Ω_0 is the initial angular velocity. Substituting the constant $\mathcal{K}(n_1)$ from Eq. (2.94) into Eq. (2.95), allows us to express the age as:

$$\tau = -\frac{1}{n_1 - 1} \frac{\Omega}{\dot{\Omega}} \left[1 - \left(\frac{\Omega}{\Omega_0} \right)^{n_1-1} \right] . \quad (2.96)$$

Assuming that the object is in a state, where $\Omega \ll \Omega_0$, and using the gravitational-wave frequency, Eq. (2.96) yields

$$\tau = -\frac{1}{n_1 - 1} \frac{f}{\dot{f}} . \quad (2.97)$$

The braking index n_1 can be measured if the second-order spin-down of the object is known, as differentiation of Eq. (2.94) gives the relation

$$n_1 = \frac{f\ddot{f}}{\dot{f}^2} . \quad (2.98)$$

Further differentiation of Eq. (2.94) yields the second braking index n_2 , also known as jerk parameter [61]. Written in terms of frequency and frequency derivatives, n_2 is

$$n_2 = \frac{f^2\ddot{\dot{f}}}{\dot{f}^3} . \quad (2.99)$$

Using Eqs. (2.97) and (2.98) we can restrict the first- and second-order spin-down bands in the range

$$-\frac{f}{\langle n_1^{\min} - 1 \rangle \tau} \leq \dot{f} \leq -\frac{f}{\langle n_1^{\max} - 1 \rangle \tau} \quad (2.100)$$

and

$$\frac{n_1^{\min} \dot{f}^2}{f} \leq \ddot{f} \leq \frac{n_1^{\max} \dot{f}^2}{f} , \quad (2.101)$$

as it has been done in a recent directed search for gravitational waves from the Cassiopeia-A (CasA) supernovae remnant [62, 63]³. Similarly to Eq. (2.101), using Eq. (2.99), we can define the search band for the third spindown, namely

$$\frac{n_2^{\min} \dot{f}^3}{f^2} \leq \ddot{\dot{f}} \leq \frac{n_2^{\max} \dot{f}^3}{f^2} . \quad (2.102)$$

³The angled brackets in Eq. (2.100) mean the averaged braking index over the lifetime of the object.

According to [61], the jerk parameter n_2 can be expressed in terms of the braking index n_1 as:

$$n_2 = 2n_1^2 - n_1 + (n_1 - 3)(1 - n_1) . \quad (2.103)$$

Note that the predicted second braking index values for the Crab pulsar, $n_2 = 10.9$, and PSR B1509-58, $n_2 = 13.6$, do not agree with the measured $n_2 = 10.23 \pm 0.03$ resp. $m = 18.3 \pm 2.9$ [61]. However, lacking other comparisons Eq. (2.103) might be an acceptable guess.

The spin-down space defined by Eqs. (2.97), (2.98) and (2.99) is model dependent through the braking indices n_1 and n_2 . These indices might be uncertain, and the shape of the parameter-space might be complicated. Therefore, it is useful to have an alternative definition for the spin-down search bands.

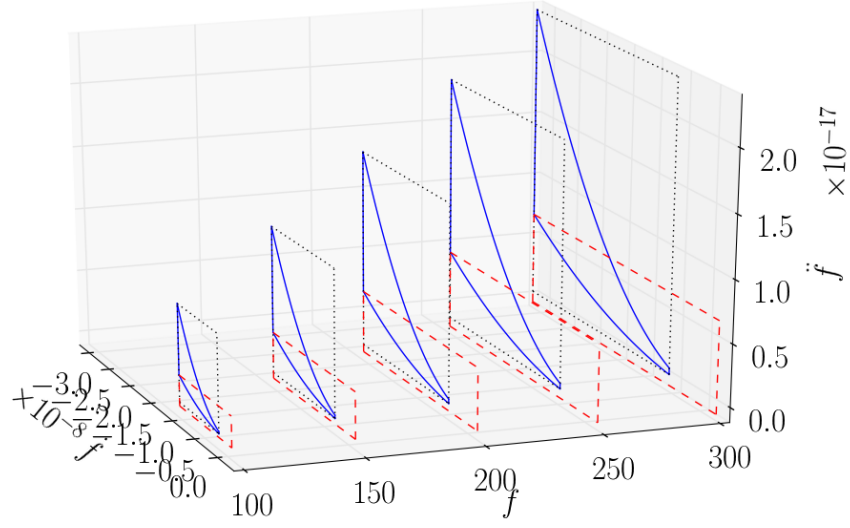
Following [64], the spindown in dimension k should be bounded by

$$|f^{(k)}| \leq k! \frac{f}{\tau^k} . \quad (2.104)$$

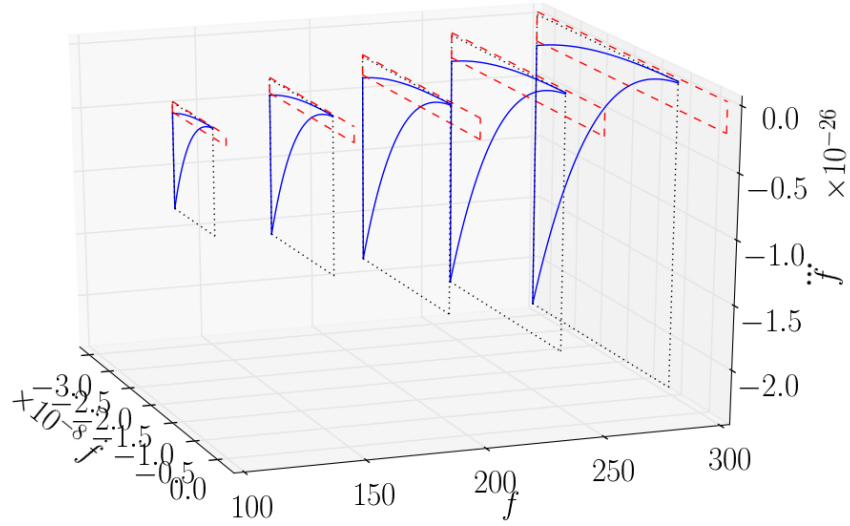
This however means that the spin-down band is frequency dependent, which may not be desirable. Therefore if we fix the spin-down range constant by fixing a minimal detectable spindown age τ_0 at some frequency f_0 , the detectable spindown age at frequency f yields

$$\tau(f) = \tau_0 f / f_0 . \quad (2.105)$$

To visualize the difference in shape between a parameter-space defined by using braking indices and by Eq. (2.104) we take as an example the parameters used to define the search space for CasA, namely $\tau = 300$ y, $n_1^{\min} = 2$, $n_1^{\max} = 7$ and $f \in [100, 300]$ Hz [63]. Using Eq. (2.103) we obtain minimal $n_2^{\min} = 7$ and maximal jerk parameter $n_2^{\max} = 67$. In Fig. 2.3(a) we plot $\ddot{f}$ as a function of f and $\dot{f}$, and in Fig. 2.3(b) we plot $\ddot{f}$ as a function of f and $\dot{f}$.



(a)



(b)

Figure 2.3.: Spindown space defined by $n_1 \in [2, 7]$ and $n_2 \in [7, 67]$ (solid blue), compared to a simple shaped orthogonal search space (red dashed). The used spindown age is $\tau = 300$ y. (a) $\ddot{f}$ as function of f and $\dot{f}$. (b) $\dddot{f}$ as function of f and $\dot{f}$

3. Optimization by Direct Search

3.1. Stochastic optimization

While analytic treatment and solution of problems in science leads to more insight in general, there are plenty of situations, where some kind of numerical optimization is unavoidable, e.g., when an analytic solution cannot be given, or the problem is untractable due to the size of the parameter space and the computational expense of a single function evaluation, which prohibits the use of a brute force numerical method. A large number of algorithms and strategies has been developed to attack such problems, e.g., Metropolis-Hastings [65], Gibbs Sampling [66], Nelder-Mead [67], Simulated Annealing [68], Nested Sampling [69], Evolutionary algorithms, MCMC, to name a few.

A major subclass of optimization algorithms are the Direct Search methods known as techniques which do not explicitly use derivatives [13]. According to [13], these techniques were introduced in the 1950s and abandoned in the early 1970s by the mathematical optimization community, mostly because they were heuristically developed, lacked proofs of convergence, and sometimes converged slowly. However these techniques remained popular, e.g., the Nelder-Mead simplex algorithm.

The numerical optimization required in Chaps. 5, 7 and 8 uses a modern incarnation of a Direct Search, namely the Mesh Adaptive Direct Search. Based on the original publications [70],[71] and [72] in the next two sections we give a short overview of the method.

3.2. Mesh Adaptive Direct Search

The Mesh Adaptive Direct Search (MADS) class of algorithms for constrained optimization is a two step - *search* and *poll* - iterative process, in which the objective is to find the minimum

$$\min_{x \in \Omega} f(x) , \quad (3.1)$$

of a function f , where $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R} \cup \infty$ is evaluated on a set of “feasible” points Ω , while derivatives are not required. Both f and Ω can be given by a “black box”, and the search is driven only by function evaluations. At iteration k the goal is to generate a trial point in Ω such that the value of the function at this point is smaller than $f(x_k)$, where x_k is the current best feasible solution.

3.2.1. The search step

The points in the *search* step may be chosen *anywhere* on the mesh [70],

$$M_k = \{x + \Delta_k^m Dz : x \in V_k, z \in \mathbb{N}^{2n}\} \subset \mathbb{R}^n, \quad (3.2)$$

where $V_k \subset \mathbb{R}^n$ is the set of all evaluated points (the cache), Δ_k^m is the mesh size parameter at iteration k , $D \equiv [I_n - I_n]$, where I_n is the n -dimensional identity matrix and z is an integer vector. For a fixed parameter $u_b > 1$, called the mesh update basis, and two given integers, namely a mesh refining exponent $w^- \leq -1$ and a mesh coarsening exponent $w^+ \geq 0$, the mesh size parameter is coarsened:

$$\Delta_{k+1}^m = u_b^{w^+} \Delta_k^m \quad (3.3)$$

if an improved mesh point is found, or refined otherwise:

$$\Delta_{k+1}^m = u_b^{w^-} \Delta_k^m. \quad (3.4)$$

The search step is optional, but may improve the overall performance of the algorithm. Therefore, different methods for selection of points on the mesh M_k have been suggested, e.g., [72] considers the usage of quadratic models.

A quadratic model m_f is a representation of a function f [72]:

$$m_f \equiv \alpha^T \phi(x) \quad (3.5)$$

defined by $q+1$ parameters $\alpha \in \mathbb{R}^{q+1}$, where $\phi(x)$ is the basis of the space of polynomials of degree up to second order in $\mathbb{R}^n$ with $q = (n+1)(n+2)/2 - 1$ elements, namely

$$\begin{aligned} \phi(x) &= \left(\phi_0(x), \phi_1(x), \dots, \phi_q(x) \right)^T \\ &= \left(1, x_1, x_2, \dots, x_n, \frac{x_1^2}{2}, \frac{x_2^2}{2}, \dots, \frac{x_n^2}{2}, x_1x_2, x_1x_3, x_{n-1}x_n \right)^T. \end{aligned} \quad (3.6)$$

To obtain α we solve the system

$$M(\phi, Y)\alpha = f(Y), \quad (3.7)$$

where $f(Y) = (y^0, y^1, \dots, y^p)^T$ and

$$M(\phi, Y) = \begin{pmatrix} \phi_0(y^0) & \phi_1(y^0) & \cdots & \phi_q(y^0) \\ \phi_0(y^1) & \phi_1(y^1) & \cdots & \phi_q(y^1) \\ \vdots & \vdots & \vdots & \vdots \\ \phi_0(y^p) & \phi_1(y^p) & \cdots & \phi_q(y^p) \end{pmatrix} \in \mathbb{R}^{p+1, q+1}.$$

If the system (3.7) is under determined or over determined, i.e., $p \neq q$, some interpolation method can be applied; see [72] for details.

To illustrate the construction of a quadratic model, consider the two-dimensional Rosenbrock function:

$$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2, \quad (3.8)$$

with minimum $f^{\min}(1, 1) = 0$. For $n = 2$ the set of basis functions is

$$\phi(x) = \left(1, x_1, x_2, \frac{x_1^2}{2}, \frac{x_2^2}{2}, x_1x_2\right)^T. \quad (3.9)$$

Therefore the quadratic model, Eq. (3.5), is given by

$$m_f = \alpha_1 + \alpha_2x_1 + \alpha_3x_2 + \alpha_4\frac{x_1^2}{2} + \alpha_5\frac{x_2^2}{2} + \alpha_6x_1x_2. \quad (3.10)$$

For simplicity we assume that the system, Eq. (3.7), is determined, i.e., we have an interpolation set Y containing six points,¹ e.g.,

$$Y = ((0.82, 0.86), (0.77, 0.75), (0.78, 0.85), (0.78, 0.87), (0.76, 0.91), (1.05, 1.11))^T, \quad (3.11)$$

for which substitution into Eq. (3.8) yields

$$f(Y) = (3.56, 2.22, 5.54, 6.20, 10.40, 0.01)^T. \quad (3.12)$$

With Eqs. (3.9), (3.11) and (3.12) the solution of Eq. (3.7) for the six coefficients α yields

$$\alpha = (39.26, -178.43, 107.55, 447.27, 197.30, -291.02)^T. \quad (3.13)$$

The two-dimensional Rosenbrock function, Eq. (3.8), and the quadratic model, Eq. (3.10), for the particular realization given through Eq. (3.13) are plotted in Fig. 3.1. The minimum of the quadratic model is found to be $m_f^{\min}(1.1, 1.08) = -0.93$.

At MADS iteration k an interpolation set Y is selected from the cache, such that all points are inside the ball $B(x_k, \rho\Delta_k^p)$, where ρ is the interpolation radius and Δ_k^p the poll size parameter, which satisfies $\Delta_k^m \leq \Delta_k^p$. The selected interpolation set is scaled, and a model is constructed. This model is optimized inside the ball $B(0, 1)$; the scaling algorithm is given in [72]. The optimal point is then rescaled back and the closest mesh point is selected. At this point the success of the search step is examined by function evaluation. If the search step is successful, the poll step is optional. The poll size parameter Δ_k^p is used to bound the poll trial points to the current incumbent solution. The exact relation between the mesh and poll size parameters will be given in the next subsection, as it depends on the method used to generate trial points in the poll step.

3.2.2. The poll step

The points generated in the poll step *must* be chosen such that they form a dense set in the unit sphere after an infinite number of iterations. This is crucial for the convergence analysis and advantage of MADS over other Direct Search methods, e.g., the Generalized Pattern Search; for details see [70].

¹All numerical values in this example are given with precision to second significant figure.

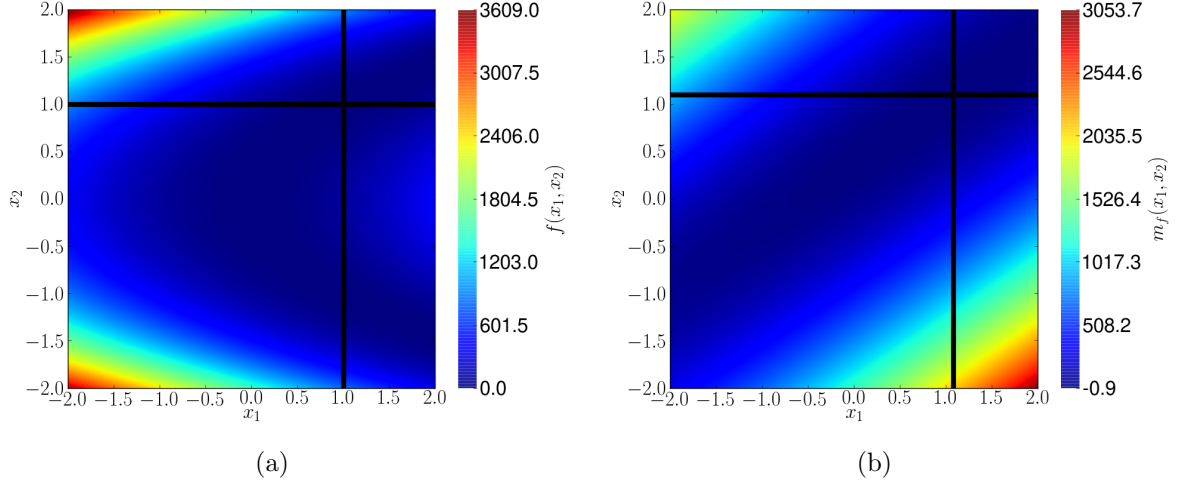


Figure 3.1.: The two-dimensional Rosenbrock function, (a), and a quadratic model of it (b). The black lines are drawn at the (x_1, x_2) pair which minimizes the function, resp. the quadratic model.

The set of poll trial points around the incumbent solution [70]

$$P_k = \{x_k + \Delta_k^m d : d \in D_k\} \subset M_k \quad (3.14)$$

is called a frame, and is called a minimal frame if the poll step fails to generate an improved solution. The optimization stops when a maximal number of iterations is exceeded or a new minimal frame cannot be found. There are two proposed methods to generate D_k , namely a stochastic and a deterministic.

The stochastic method introduced in [70], called LtMads, involves the generation of a random integer nonsingular matrix B_k , which is completed either by $-B_k$, or a vector $b(l_k)$, to form a minimal positive basis $D_k = [B_k \ b(l_k)]$ or a maximal basis $D_k = [B_k \ -B_k]$. The integer l_k is related to the mesh size ²

The construction of B_k is done as follows. First, the vector $b(l_k)$ is constructed such, that one random element $\hat{i}$ is either -2^{l_k} or 2^{l_k} and the other elements are integers between $-2^{l_k} + 1$ and $2^{l_k} - 1$. For example, with $n = 4, l_0 = 0$, one possible realization of b_0 is $b_0 = (0, -1, 0, 0)^T$. Second, a $(n - 1) \times (n - 1)$ lower triangular matrix L is generated, where each diagonal element is $\pm 2^{l_k}$ and the other components are randomly chosen in $\{-2^{l_k} + 1, \dots, 2^{l_k} - 1\}$; let

$$L = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Third, a row of zeros is inserted into L at position $\hat{i}$ and $b(l_k)$ is appended, then the columns of L undergo random permutation, resulting in a basis B_k in $\mathbb{R}^n$. A possible

²With $l_{k=0} = 0$, if the iteration is successful $l_{k+1} = l_k + 1$, otherwise $l_{k+1} = l_k - 1$ [71].

basis obtained from the given L is

$$B_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

which yields

$$D_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \end{pmatrix}$$

When the minimal basis is used, $\Delta_k^p = n\sqrt{\Delta_k^m}$, otherwise $\Delta_k^p = \sqrt{\Delta_k^m}$. A two-dimensional example of MADS frames is given in Fig. 3.2.

An improved method for the generation of D_k has been proposed in [71], where pseudo-random Halton sequences are used to construct an orthogonal integer basis H_k and with this $D_k = [H_k - H_k]$. The construction of H_k is done in three steps. First, a pseudo-random Halton sequence is used to generate a vector in $[0, 1]^n$. The t -th element of a Halton sequence (Halton directions) is given by [71]

$$u_t = (u_{t,p_1}, u_{t,p_2}, \dots, u_{t,p_n})^T \in [0, 1]^n, \quad (3.15)$$

where p_j is the j -th prime number and $u_{t,p}$ is the radical-inverse function in basis p , namely

$$u_{t,p} = \sum_{r=0}^{\infty} \frac{a_{t,r,p}}{p^{1+r}}, \quad (3.16)$$

with $a_{t,r,p} \in \mathbb{Z}_+$ the unique coefficients of the base p expansion of t (Halton index):

$$t = \sum_{r=0}^{\infty} a_{t,r,p} p^r. \quad (3.17)$$

For example, with $n = 4$, the first four prime numbers are $p \equiv \{2, 3, 5, 7\}$ and for $t = 7$ we obtain $u_7 = (7/8, 5/9, 11/25, 1/49)^T$. Second, the Halton directions are adjusted to directions with norm close to $2^{|l_k|/2}$, namely [71]

$$q_t(\alpha) = \text{round} \left(\alpha \frac{2u_t - e}{\|2u_t - e\|} \right) \in \mathbb{Z}^n \cap \left[-\alpha - \frac{1}{2}, \alpha + \frac{1}{2} \right]^n, \quad (3.18)$$

where $\alpha \in \max_{\alpha \in \mathbb{R}_+} \|q_t(\alpha)\|$ is a scaling factor with $\|q_t(\alpha)\| \leq 2^{|l_k|/2}$. For $l_0 = 0$, u_7 is adjusted to $q_7 = (0, 0, 0, 1)^T$. Third, to produce n orthogonal and integer vectors, the symmetric scaled Housholder transformation H is applied to q_t , i.e., [71]

$$H = \|q\|^2 (I_n - 2vv^T), \quad (3.19)$$

where $v = q/\|q\|$. For example, by plugging q_7 into Eq. (3.19) we obtain

$$H_0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix},$$

and therefore

$$D_0 = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The advantage of OrthoMads over LtMads is not only its deterministic nature, but also the better covering of the parameter space. For this method of generation of D_k , the relation between the mesh size Δ_k^m and the poll size parameter Δ_k^p is given by $\Delta_k^p = 2^{|l_k|} \Delta_k^m$.

To demonstrate the performance of the LtMads and OrthoMads methods for generation of poll points, we consider the two-dimensional Rosenbrock function given in Eq. 3.8. For simplicity and better comparison we omit the search step in this example. The result of the optimization is plotted in Fig. 3.3. Using the same initial starting point $(0, 0)$ and boundaries $x_1 \in [-2, 2]$ and $x_2 \in [-2, 2]$ the OrthoMads method achieves a better solution in terms of absolute value and position, namely $f_O^{\min}(0.92, 0.85) = 0.005$, compared to the optimal solution obtained with the LtMads method, namely $f_L^{\min}(0.83, 0.69) = 0.02$. For both methods the maximal number of iterations is limited to 200.

3.2.3. NOMAD

A reference implementation of MADS, called NOMAD, written in the C++ programming language [73] is provided by the authors of the MADS algorithm under the GNU Lesser General Public License. The version 3.5.1 of this software is used for the optimization required in Chaps. 5, 7 and 8³. The NOMAD library is feature rich, however listing all control switches and parameters here is out of scope. Thus, if the value of a control parameter is not explicitly stated where appropriate, the default value as of the given software version is used.

³The examples in this chapter are generated using my own rudimentary MADS implementation written in the Python programming language.

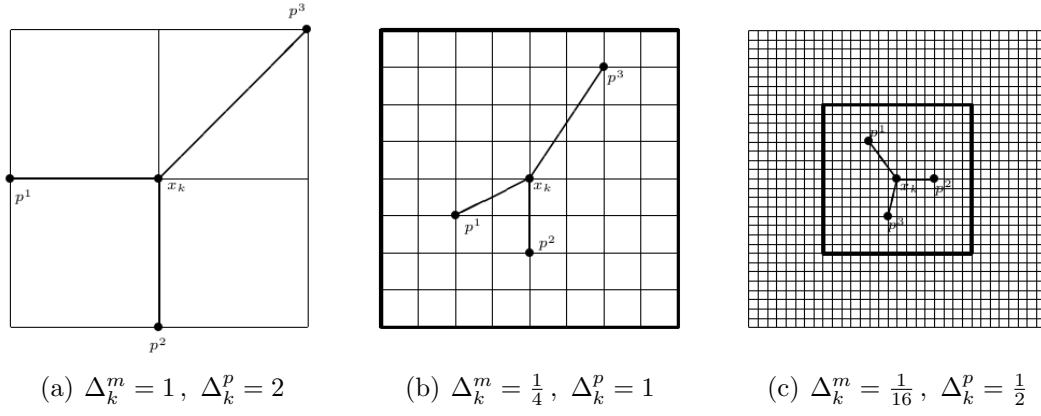


Figure 3.2.: A two-dimensional example, $n = 2$, of MADS frames $P_k = \{x_k + \Delta_k^m d : d \in D_k\} = \{p^1, p^2, p^3\}$, with $d_{1,2} \in \{-1, 0, 1\}$ for different values of Δ_k^m and Δ_k^p , with $\Delta_k^p = n\sqrt{\Delta_k^m}$. The mesh M_k is the intersection of all lines.[70].

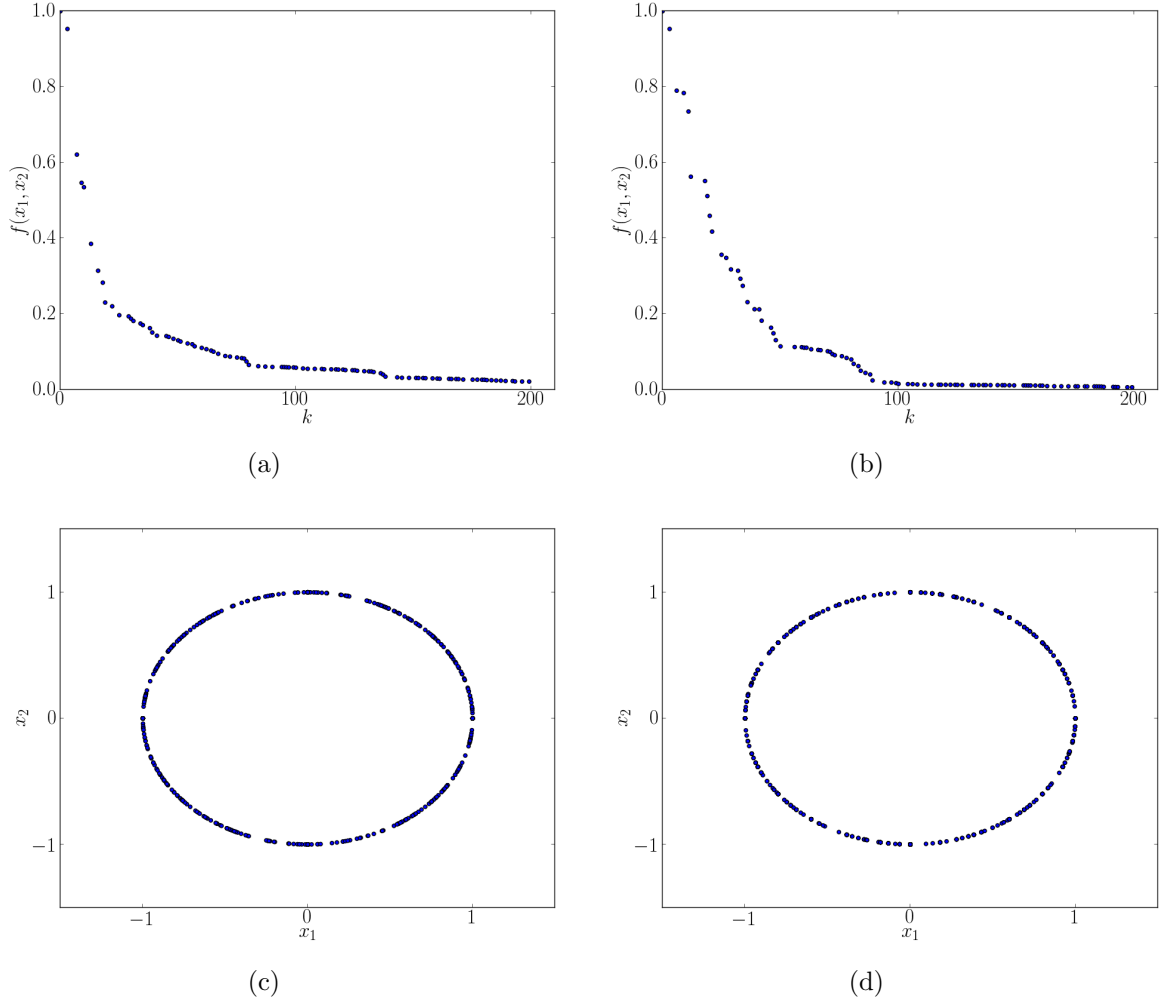


Figure 3.3.: LtMads and OrthoMads optimization of the two-dimensional Rosenbrock function. (a) Improvement of the objective as a function of iteration k using LtMads. (b) Improvement of the objective using OrthoMads. (c) Normalized poll directions using LtMads. (d) Normalized poll directions using OrthoMads.

Part II.

Semicoherent Search Optimization

Summary

Coherent wide parameter-space searches for continuous gravitational waves are typically limited in sensitivity by their prohibitive computing cost. Therefore semicoherent methods (such as StackSlide) can often achieve a better sensitivity.

In Chap. 4 we develop an analytical method for finding *optimal* StackSlide parameters at fixed computing cost under ideal conditions of gapless data with Gaussian stationary noise. This solution separates two regimes: an *unbounded* regime, where it is always optimal to use all the data, and a *bounded* regime with a finite optimal observation time. The analysis of the sensitivity scaling reveals that both the fine- and coarse-grid mismatches contribute equally to the average StackSlide mismatch, an effect that had been overlooked in previous studies. We discuss various practical examples for the application of this optimization framework, illustrating the potential gains in sensitivity compared to previous searches.

In Chap. 5 we extend the analytical framework developed in Chap. 4 to take into account the possible gaps in the real detector data and the noise level changes. We show, how the sensitivity optimization can be decoupled from the data selection problem. To find optimal semicoherent search parameters we therefore apply a numerical optimization using as example the semicoherent StackSlide search. We also describe a suitable data selection algorithm. Thus the outcome of the numerical optimization are the optimal search parameters and the selected data, which is needed to perform the search in practice. We first test the numerical optimization procedure under ideal conditions and obtain the results of the analytical method. Then we gradually relax the requirements on the data to finally use real detector data.

4. Optimal StackSlide method at fixed computing cost

Motivation. The detection of continuous gravitational waves (CWs) from spinning neutron stars (NSs) in our galaxy remains an elusive goal, despite the global network of detectors LIGO [74], Virgo [75] and GEO 600 [76] having completed their initial and enhanced science runs (e.g. see [77, 78, 79, 63]). The search for CWs will likely remain a difficult challenge with uncertain prospects even in the era of Advanced detectors [80, 81, 82] and third-generation detectors such as ET [83]. Two main reasons for this are (i) astrophysical priors on CWs and (ii) the large parameter space of signal parameters to explore (cf. [11] for a review and further references).

(i) Current astrophysical priors contain large uncertainties on the expected strength of CW emissions from spinning NSs, with a strong bias towards extremely weak signals, informed by the population of known pulsars as well as by a statistical analysis of a putative galactic “gravitar” population [84]. (ii) The required number of templates for a coherent matched-filter search over a range of unknown signal parameters typically grows very rapidly with increasing duration of data analyzed. Therefore only a fraction of the available data can be analyzed coherently (e.g. see [53, 62, 85]).

It was realized early on [64] that in situations where the total computing cost of the search is constrained, a *semicoherent* approach could typically achieve better sensitivity than coherent matched filtering: shorter segments of data are analyzed coherently, then the statistics from these segments are combined incoherently. One method of incoherent combination simply consists of summing the statistics from the different segments, which is typically referred to as the “StackSlide” method in the CW context (also known as the Radon transform). The template bank used for the semicoherent combination of coherent statistics is referred to as the *fine grid*, as it typically requires a higher resolution than the template banks of the per-segment coherent searches (referred to as *coarse grids*). Details of the respective template banks will be discussed in Sec. 4.2.4.

There are a number of different semicoherent methods: for example, recent work [86] has shown that StackSlide sensitivity can be improved by a sliding coherence-window approach. A closely related variant to StackSlide is the *Hough transform* [43], which counts the number of segments in which the statistic crosses a given threshold, instead of summing the statistics. This is generally less sensitive, but is designed to be more robust in the presence of strong non-stationarities. A somewhat different semicoherent approach are cross-correlation methods, described in more detail in [87].

Related to the semicoherent methods are the so-called *hierarchical* schemes, which

consist of following up “promising” candidates from a (coherent or semicoherent) search by subsequent, more sensitive searches, referred to as “stages”. This procedure is iterated until the parameter space of surviving candidates is sufficiently narrowed down for a fully coherent follow-up using all the data. Work on implementing such schemes in practice is still ongoing. In Chap. 7 we discuss a possible two-stage approach.

Optimization problem. In this chapter we focus on the standard single-stage StackSlide method, which was also used in previous optimization studies [64, 88], and is relatively straightforward to describe analytically.

Any search method contains a number of tuneable parameters, such as the template-bank mismatch, the data selection procedure, and the number and length of segments to analyze. Hierarchical schemes would further require specification of the number of stages and the respective distributions of computing costs and candidate thresholds. The sensitivity of a search generally depends on all these choices, and we therefore need to study how to maximize sensitivity as a function of these parameters.

This optimization problem has been studied previously by Brady and Creighton [64] (henceforth ‘BC’) and subsequently by Cutler, Gholami and Krishnan [88] (in the following ‘CGK’). Both studies have focused on the wider problem of optimizing a multi-stage hierarchical scheme of StackSlide stages, and have directly resorted to fully numerical exploration of the optimization problem. Here instead we focus on the simpler single-stage search, which allows us to fully analytically analyze the problem. In the next step this can be used as a building block to attack the optimization of hierarchical schemes.

Note that for a network of detectors with different noise-floors, the choice of detectors to use at fixed computing cost is part of the optimization problem, but under the present assumption of “ideal data” the answer can be obtained independently [89]. More work is required to develop a practical algorithm to compute the optimal search parameters for given data from a network of detectors, including gaps, non-stationarities and various detector artifacts. In the next chapter we discuss a step toward solution of this problem.

Summary of main results. Careful analysis of the sensitivity scaling shows that the average StackSlide mismatch is given by the *sum* of the average mismatches from the coarse- and fine-grid template banks, an effect that had previously been overlooked. Note that we allow for independent coarse- and fine-grid mismatches, while BC and CGK forced them to be equal as an *ad-hoc* constraint.

The analytic optimization is achieved by using local power-law approximations to the computing-cost and sensitivity functions. The results provide analytic self-consistency conditions for the optimal solution: if the initial power-law coefficients agree with those found at the analytic solution, then the solution is self-consistent and (locally) optimal. If this is not the case, one can iterate over successive solutions or scan a range of StackSlide parameters, in order to “bootstrap” into a self-consistent optimal solution.

We find that the analytic solution for StackSlide searches separates two different regimes depending on the power-law coefficients: a *bounded* regime in which there is a finite optimal observation time, and an *unbounded* regime in which the optimal solution always consists of using all of the available data, irrespective of the available

computing-cost.

This chapter is organized as follows. In Sec. 4.1 we discuss the general CW optimization problem, which includes the single-stage StackSlide search as the lowest-level building block. In Sec. 4.2 we derive the sensitivity estimate and computing-cost functions for StackSlide searches, and motivate their approximation as local power-laws. After deriving in Sec. 4.3 the general analytical solution and discussing a few special cases, we provide examples for the practical application of this framework in Sec. 4.4: directed searches, all-sky searches, and searches for CWs from NSs in binary systems. We discuss in Sec. 4.5.

4.1. Maximizing probability of a CW detection

The goal for wide parameter-space CW searches for unknown signals should be to maximize the probability of detection, given current astrophysical priors, detector data, and finite computing resources. Conceptually one can think of this problem as a hierarchy of two questions:

- (i) What parameter-space $\mathbb{P} \subseteq \mathbb{P}^{(0)}$ to search? More generally: how to distribute the total available computing power C_0 over the space $\mathbb{P}^{(0)}$ of possible CW signals, given astrophysical priors, detector data and an (optimal) search method?
- (ii) What is the optimal search method? Namely, which method yields the highest detection probability p_{det} on a parameter-space cell $\Delta\mathbb{P}$ if we spend computing-cost ΔC on it?

The answer to the first question relies on the second, but we can analyze the lower-level second question independently of the first. There has been surprisingly little work on this problem so far. The first question has hardly been addressed at all, except for recent work by Knispel [90]. The second question has been studied previously by BC [64] and CGK [88], assuming a specific type of hierarchical scheme, which we refer to as the *classical hierarchical scheme* (CHS).

In the CHS one performs a hierarchy of semicoherent searches (called *stages*), starting with a relatively low-sensitivity search over the whole initial parameter space $\mathbb{P}^{(1)}$. Promising candidates crossing the first-stage threshold are selected and constitute the search subspace $\mathbb{P}^{(2)} \subset \mathbb{P}^{(1)}$ for the second, higher-sensitivity stage. This is iterated until eventually after m such stages a fully coherent search over all the data can be performed on the surviving candidates. At this point one has reached the maximal possible sensitivity for a small portion $\mathbb{P}^{(m)} \subset \mathbb{P}^{(1)}$ of the initial parameter space.

Each stage (i) is characterized by its input parameter-space $\mathbb{P}^{(i)}$, a computing-cost constraint $C_0^{(i)}$ and a false-alarm probability $p_{\text{fA}}^{(i)}$. Each stage selects a candidate subspace $\mathbb{P}^{(i+1)} \subset \mathbb{P}^{(i)}$ to follow up in the next stage. An optimal per-stage search would result in the highest detection probability $p_{\text{det}}^{(i)}$ for given signal strength h_{rms} and constraints $\{p_{\text{fA}}^{(i)}, \mathbb{P}^{(i)}, C_0^{(i)}\}$. The tuneable CHS parameters are therefore the number m of stages

and the per-stage constraints $\{p_{\text{fA}}^{(i)}, C_0^{(i)}\}$. These can be varied in order to maximize the overall detection probability $p_{\text{det}}(h_{\text{rms}})$ for the given total signal parameter-space $\mathbb{P}^{(0)}$, computing cost $C_0 = \sum_{i=1}^m C_0^{(i)}$ and false-alarm probability $p_{\text{fA}} = \prod_{i=1}^m p_{\text{fA}}^{(i)}$.

This formulation of the CHS suggests that each stage (i) could be considered an independent optimization problem for given external constraints $\{p_{\text{fA}}^{(i)}, \mathbb{P}^{(i)}, C_0^{(i)}\}$, if none of its internal parameters interfere with the overall hierarchical scheme. One might contend that the parameter-space resolution of the search violates this clean factorization: the follow-up space $\mathbb{P}^{(i+1)}$ from stage (i) depends on its parameter-space resolution, which might impact the required computing-cost of the next stage. However, it is easy to see that (to first order) such a coupling is not expected. The number of candidates $\mathcal{N}_{\text{cand}}$ returned from any stage (except for the last one) will be dominated by the number $\mathcal{N}_{\text{fA}}$ of false alarms. Therefore $\mathcal{N}_{\text{cand}} \approx \mathcal{N}_{\text{fA}} \approx p_{\text{fA}} \mathcal{N}$, where $\mathcal{N}$ is the number of (approximately) independent templates searched in this stage. This can be estimated as $\mathcal{N} \approx V_{\mathbb{P}}/v_0$, in terms of the (metric) volume $V_{\mathbb{P}}$ of the parameter space $\mathbb{P}$, and the volume v_0 covered by one template. Therefore the *number* $\mathcal{N}_{\text{cand}}$ of follow-up candidates from any stage does indeed depend on its parameter-space resolution, which depends on the internal stage parameters. However, the computing cost of the next stage depends primarily on the *volume* of the follow-up parameter-space, which is $V_{\text{fA}} \approx \mathcal{N}_{\text{fA}} v_0 \approx p_{\text{fA}} V_{\mathbb{P}}$, and is therefore *independent* of internal stage parameters. It is interesting to note that each stage (i) in this scheme achieves a reduction of the input parameter-space volume by roughly a factor of the false-alarm probability $p_{\text{fA}}^{(i)}$, irrespective of the internal details of that search.

The optimal per-stage search method is essentially unknown, but following BC and CGK we focus on a known good strategy, namely the StackSlide method. While different semicoherent methods differ in the details and their exact sensitivity, they share the main characteristics of coherently searching N shorter segments of length ΔT , and combining them incoherently in some way. We roughly expect the *sensitivity per cost* of different methods to behave qualitatively similarly to the StackSlide method, but more work would be required to study this in detail.

4.2. Properties of a single-stage StackSlide search

The general StackSlide scheme consists of dividing the data (of total duration T) into N segments of duration $\Delta T = T/N$, then performing a coherent matched-filter search on each segment and combining these statistics *incoherently* to a new statistic Σ by summing them across segments. The coherent matched-filter statistic used is the $\mathcal{F}$ -statistic, which was first derived in [91] and extended to multiple detectors in [46]. Using the same amount of data as a fully coherent search, the resulting semicoherent statistic is less sensitive, but substantially cheaper to compute over a wide parameter space. At *fixed computing cost* a semicoherent search is therefore generally more sensitive than a fully coherent $\mathcal{F}$ -statistic search.

Notation: we distinguish quantities Q that can refer to either the coherent or the in-

coherent step in the following way: we use a tilde, i.e. $\tilde{Q}$ when referring to the coherent step, and a hat, i.e. $\hat{Q}$ when referring to the incoherent step. For the following derivations we restrict ourselves to a single-detector formalism for simplicity, but we state the (trivial) generalization to $N_{\text{det}} \geq 1$ detectors of relevant results.

4.2.1. The StackSlide search method

Let $k = 1 \dots N$ be the index over segments, and $\lambda \in \mathbb{P}$ a point in the search space $\mathbb{P}$ of signal parameters. The “ideal” StackSlide statistic $\Sigma_0(\lambda)$ is defined as

$$\Sigma_0(\lambda) \equiv \sum_{k=1}^N 2\mathcal{F}_k(\lambda), \quad (4.1)$$

i.e. a simple sum of $\mathcal{F}$ -statistic values $\{2\mathcal{F}_k(\lambda)\}_{k=1}^N$ computed at *the same* template point λ across all N segments.

This would require computing the $\mathcal{F}$ -statistic over the same template bank as Σ_0 in every segment. However, the metric resolution of Σ_0 is generally finer than that of the single-segment $\mathcal{F}$ -statistics [56], and therefore this approach would spend unnecessary computing cost on the coherent $\mathcal{F}$ -statistic. In practice $\mathcal{F}$ is therefore computed over a *coarse grid* of $\tilde{N}$ templates in each segment k , and is interpolated in order to sum $\mathcal{F}$ on the *fine grid* of $\hat{N} \geq \tilde{N}$ templates (e.g. see [92]).

Typically the interpolation consists of picking the *closest* (in the metric sense) coarse-grid point $\tilde{\lambda}_k(\hat{\lambda})$ to the fine-grid point $\hat{\lambda}$ from every segment k , i.e. we approximate Eq. (4.1) by

$$\Sigma(\hat{\lambda}) \equiv \sum_{k=1}^N 2\mathcal{F}_k(\tilde{\lambda}_k(\hat{\lambda})) \approx \Sigma_0(\hat{\lambda}), \quad (4.2)$$

which we refer to as the “interpolating” StackSlide statistic Σ . The following sensitivity optimization focuses exclusively on this interpolating StackSlide method, which is the most relevant approach for current practical applications. The subtle difference between *interpolating StackSlide* Σ and *ideal StackSlide* Σ_0 with respect to its sensitivity and mismatches has been overlooked in previous studies, and will be important for the optimization problem.

4.2.2. Mismatch and metric

$\mathcal{F}$ -statistic mismatch

In the presence of a signal timeseries $s(t, \lambda_s)$ with phase parameters λ_s , the statistic $2\mathcal{F}_k(\tilde{\lambda})$ in a point $\tilde{\lambda}$ follows a non-central χ^2 -distribution with four degrees of freedom and non-centrality parameter $\rho_k^2(\lambda_s, \tilde{\lambda})$. We denote this probability distribution as

$$P(2\mathcal{F}_k | \rho_k^2) = \chi_4^2(2\mathcal{F}_k; \rho_k^2), \quad (4.3)$$

which has the expectation value

$$E \left[2\mathcal{F}_k(\lambda_s, \tilde{\lambda}) \right] = 4 + \rho_k^2(\lambda_s, \tilde{\lambda}). \quad (4.4)$$

The quantity ρ_k is often referred to as the coherent signal-to-noise ratio (SNR). In the case of a perfectly-matched template $\tilde{\lambda} = \lambda_s$, the resulting “optimal” SNR [91] in segment k can be expressed as

$$\begin{aligned} \rho_k^2(\lambda_s, \lambda_s) &= \frac{2}{S_n} \int_{t_k}^{t_k + \Delta T} s^2(t, \lambda_s) dt \\ &\equiv \frac{2}{S_n} h_{\text{rms},k}^2 \Delta T, \end{aligned} \quad (4.5)$$

where t_k is the start-time of the k th segment, S_n is the (single-sided) noise power spectral density at the signal frequency f_s . In the second equality we defined the rms signal strength $h_{\text{rms},k}$ [88] in segment k , which is a useful measure of the *intrinsic* signal strength in the detectors, independently of the quality and the amount of data used.

The signal strength $h_{\text{rms},k}$ depends on the intrinsic signal amplitude h_0 , the sky-position, polarization angles, and detector orientation during segment k . One can show [91, 44] that averaging $h_{\text{rms},k}^2$ isotropically over sky-positions and polarization angles yields the relation $\langle h_{\text{rms},k}^2 \rangle_{\text{sky,pol}} = (2/25) h_0^2$. Furthermore, for segment lengths of order $\Delta T \gtrsim \mathcal{O}(1 \text{ days})$, the averaging in Eq. (4.5) results in $h_{\text{rms},k}$ tending towards a constant over all segments. Therefore it will be convenient to approximate $h_{\text{rms},k} \approx h_{\text{rms}}$, and so we can write

$$\rho_k^2(\lambda_s, \lambda_s) \approx \frac{2}{S_n} h_{\text{rms}}^2 \Delta T \equiv \rho_{\text{opt}}^2(\Delta T), \quad (4.6)$$

which defines the average optimal SNR ρ_{opt} for given segment length ΔT .

Note that this approximation only applies to the perfectly-matched SNR $\rho_k(\lambda_s, \lambda_s)$. The “mismatched” SNR $\rho_k(\lambda_s, \tilde{\lambda})$ in an offset template $\tilde{\lambda} = \lambda_s + \Delta\lambda$ is reduced with respect to the optimal SNR $\rho_{\text{opt}}(\Delta T)$. The corresponding relative loss defines the (segment-specific) *mismatch* function $\tilde{\mu}_k(\lambda_s, \tilde{\lambda})$, namely

$$\begin{aligned} \tilde{\mu}_k(\lambda_s, \tilde{\lambda}) &\equiv 1 - \frac{\rho_k^2(\lambda_s, \tilde{\lambda})}{\rho_{\text{opt}}^2(\Delta T)} \\ &= \tilde{g}_{ij,k}(\lambda_s) \Delta\lambda^i \Delta\lambda^j + \mathcal{O}(\Delta\lambda^3), \end{aligned} \quad (4.7)$$

where Taylor-expansion for small offsets $\Delta\lambda$ defines the (coherent) metric tensor $\tilde{g}_{ij,k}(\lambda)$ for segment k . The concept of the parameter-space metric was first introduced in [52, 51], and analyzed in the context of a simplified CW statistic [53] and the $\mathcal{F}$ -statistic [41].

The per-segment *coarse-grid* template bank is constructed under the constraint that no signal point $\lambda_s \in \mathbb{P}$ should exceed a given *maximal mismatch* $\tilde{m}$ to its closest (i.e. with the smallest mismatch) coarse-grid template $\tilde{\lambda}_k(\lambda_s)$, namely

$$\tilde{\mu}_k(\lambda_s, \tilde{\lambda}_k(\lambda_s)) \leq \tilde{m} \quad \text{for all } \lambda_s \in \mathbb{P}. \quad (4.8)$$

Mismatch of “ideal” StackSlide

The “ideal” StackSlide statistic Σ_0 defined in Eq. (4.1) is the basis for the definition of the semicoherent metric [64, 56, 93]). The statistic Σ_0 follows a non-central χ^2 -distribution with $4N$ degrees of freedom, denoted as

$$P(\Sigma_0 | \rho_{\Sigma_0}^2) = \chi_{4N}^2(\Sigma_0; \rho_{\Sigma_0}^2), \quad (4.9)$$

with non-centrality parameter

$$\rho_{\Sigma_0}^2(\lambda_s, \hat{\lambda}) \equiv \sum_{k=1}^N \rho_k^2(\lambda_s, \hat{\lambda}), \quad (4.10)$$

where λ_s are the signal parameters and $\hat{\lambda}$ is the location of a fine-grid template. The corresponding expectation value is

$$E[\Sigma_0(\lambda_s, \hat{\lambda})] = 4N + \rho_{\Sigma_0}^2. \quad (4.11)$$

The perfectly-matched non-centrality parameter $\rho_{\Sigma_0}^2(\lambda_s, \lambda_s)$ can be expressed as

$$\begin{aligned} \rho_{\Sigma_0}^2(\lambda_s, \lambda_s) &= \sum_{k=1}^N \rho_k^2(\lambda_s, \lambda_s) \\ &= N \rho_{\text{opt}}^2(\Delta T) \\ &= \rho_{\text{opt}}^2(T), \end{aligned} \quad (4.12)$$

which is identical to that of a perfectly-matched $\mathcal{F}$ -statistic over the same total duration T , as seen from Eq. (4.6). The reason why the StackSlide statistic Σ_0 is less sensitive than the $\mathcal{F}$ -statistic for the same amount of data T stems from the different degrees of freedom of the respective distributions, namely $\chi_4^2(\rho^2)$ for the $\mathcal{F}$ -statistic as opposed to $\chi_{4N}^2(\rho^2)$ for StackSlide Σ_0 .

The mismatch function $\hat{\mu}_0(\lambda_s, \hat{\lambda})$ of ideal StackSlide is defined in analogy to Eq. (4.7) as

$$\begin{aligned} \hat{\mu}_0(\lambda_s, \hat{\lambda}) &\equiv 1 - \frac{\rho_{\Sigma_0}^2(\lambda_s, \hat{\lambda})}{\rho_{\text{opt}}^2(T)} \\ &= \hat{g}_{ij}(\lambda_s) \Delta \lambda^i \Delta \lambda^j + \mathcal{O}(\Delta \lambda^3), \end{aligned} \quad (4.13)$$

where $\Delta \lambda \equiv \hat{\lambda} - \lambda_s$ is the offset between the fine-grid template $\hat{\lambda}$ and the signal location λ_s , and Taylor-expansion in small $\Delta \lambda$ defines the (semicoherent) metric tensor $\hat{g}_{ij}$. Using Eqs. (4.12) and (4.7), we can rearrange the expression for the mismatch as

$$\begin{aligned} \hat{\mu}_0(\lambda_s, \hat{\lambda}) &= \frac{1}{N} \sum_{k=1}^N \tilde{\mu}_k(\lambda_s, \hat{\lambda}) \\ &\approx \left(\frac{1}{N} \sum_{k=1}^N \tilde{g}_{ij,k}(\lambda_s) \right) \Delta \lambda^i \Delta \lambda^j, \end{aligned} \quad (4.14)$$

which shows that the ideal StackSlide mismatch $\hat{\mu}_0$ and metric $\hat{g}_{ij}$ are segment-averages of the coherent mismatches and metrics, respectively.

The *fine-grid* template bank of a StackSlide search is constructed under the constraint that no signal point $\lambda_s \in \mathbb{P}$ should exceed a given *maximal mismatch* $\hat{m}$ to its *closest* (i.e. with the smallest mismatch) fine-grid template $\hat{\lambda}(\lambda_s)$, namely

$$\hat{\mu}_0(\lambda_s, \hat{\lambda}(\lambda_s)) \leq \hat{m} \quad \text{for all } \lambda_s \in \mathbb{P}. \quad (4.15)$$

Mismatch of “interpolating” StackSlide

We can now combine the above results to derive the mismatch of the interpolating StackSlide statistic Σ defined in Eq. (4.2). This statistic follows a non-central χ^2_{4N} distribution, namely

$$P(\Sigma | \rho_\Sigma^2) = \chi^2_{4N}(\Sigma; \rho_\Sigma^2), \quad (4.16)$$

with non-centrality parameter

$$\rho_\Sigma^2(\lambda_s, \hat{\lambda}) \equiv \sum_{k=1}^N \rho_k^2(\lambda_s, \tilde{\lambda}_k(\hat{\lambda})), \quad (4.17)$$

where λ_s are the signal phase parameters, and $\tilde{\lambda}_k(\hat{\lambda})$ is the closest coarse-grid template in segment k to the fine-grid point $\hat{\lambda}$.

The mismatch function $\hat{\mu}(\lambda_s, \hat{\lambda})$ of interpolating StackSlide is therefore

$$\begin{aligned} \hat{\mu}(\lambda_s, \hat{\lambda}) &\equiv 1 - \frac{\rho_\Sigma^2(\lambda_s, \hat{\lambda})}{\rho_{\text{opt}}^2(T)} \\ &= \frac{1}{N} \sum_{k=1}^N \tilde{\mu}_k(\lambda_s, \tilde{\lambda}_k(\hat{\lambda})), \end{aligned} \quad (4.18)$$

which allows us to express the mismatched non-centrality parameter as

$$\rho_\Sigma^2(\lambda_s, \hat{\lambda}) = \left(1 - \hat{\mu}(\lambda_s, \hat{\lambda})\right) \rho_{\text{opt}}^2(T). \quad (4.19)$$

The extra offset per-segment, $\delta\lambda_k \equiv \tilde{\lambda}_k(\hat{\lambda}) - \hat{\lambda}$, incurred due to using the closest coarse-grid point $\tilde{\lambda}_k(\hat{\lambda})$ instead of the fine-grid point $\hat{\lambda}$ tends to increase the mismatch with respect to the ideal mismatch function $\hat{\mu}_0$ of Eq. (4.14). In order to quantify this effect, we write the effective per-segment offset from a signal as $\Delta\tilde{\lambda}_k \equiv \tilde{\lambda}_k(\hat{\lambda}) - \lambda_s$, while the ideal per-segment offset would be $\Delta\hat{\lambda} \equiv \hat{\lambda} - \lambda_s$. We can write $\Delta\tilde{\lambda}_k = \Delta\hat{\lambda} + \delta\lambda_k$, and inserting this into the coherent-metric of Eq. (4.7) we obtain (neglecting higher-order terms $\mathcal{O}(\Delta\lambda^3)$):

$$\begin{aligned} \tilde{\mu}_k(\lambda_s, \tilde{\lambda}_k(\hat{\lambda})) &= \tilde{g}_{ij,k} \Delta\tilde{\lambda}_k^i \Delta\tilde{\lambda}_k^j \\ &= \tilde{\mu}_k(\lambda_s, \hat{\lambda}) + \tilde{g}_{ij,k} \delta\lambda_k^i \delta\lambda_k^j + 2\tilde{g}_{ij,k} \Delta\hat{\lambda}^i \delta\lambda_k^j, \end{aligned} \quad (4.20)$$

where in the first term we recover the ideal per-segment mismatch function of Eq. (4.14), the second term represents an extra mismatch due to the offset $\delta\lambda_k$, while the last term depends on the opening angle θ_k of the mismatch triangle (see Fig. 4.1), namely $2|\Delta\hat{\lambda}||\delta\lambda_k|\cos\theta_k$, with mismatch norm defined as $|x|^2 \equiv \tilde{g}_{ij,k}x^ix^j$.

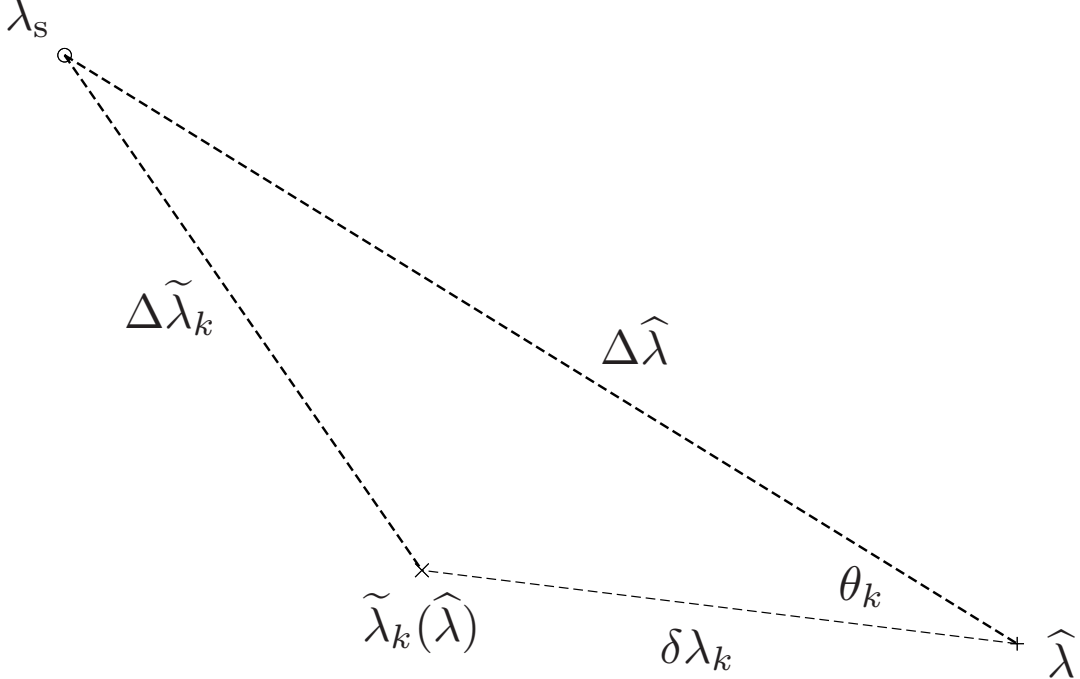


Figure 4.1.: Mismatch triangle formed by the signal point λ_s , closest fine-grid template $\hat{\lambda}$, and the coarse-grid template $\tilde{\lambda}_k(\hat{\lambda})$ closest to $\hat{\lambda}$ in segment k .

We assume that the fine-grid point $\hat{\lambda}$ falls randomly into the Wigner-Seitz cell of the closest coarse-grid template $\tilde{\lambda}_k(\hat{\lambda})$ in segment k . Given that the coarse-grid metric $\tilde{g}_{ij,k}$ generally varies across segments, we further assume that the offset $\delta\lambda_k$ approximates a uniform random sampling of the coarse-grid Wigner-Seitz cell. Inserting Eq. (4.20) into (4.18), we see that the average over the angle-term $|\Delta\hat{\lambda}||\delta\lambda_k|\cos\theta_k$ will tend to zero, as any sign of $\cos\theta_k$ is equally likely, while the average norm $|\delta\lambda_k|^2$ will tend to the average mismatch $\langle\tilde{\mu}\rangle$ of the coarse-grid template bank, and so we obtain

$$\hat{\mu}(\lambda_s, \hat{\lambda}) \approx \hat{\mu}_0(\lambda_s, \hat{\lambda}) + \langle\tilde{\mu}\rangle . \quad (4.21)$$

When estimating the sensitivity of the interpolating StackSlide statistic, we will further average this expression over randomly-chosen signal locations λ_s , and therefore the above approximate averaging expressions will become exact, and we obtain

$$\langle\hat{\mu}\rangle = \langle\hat{\mu}_0\rangle + \langle\tilde{\mu}\rangle , \quad (4.22)$$

where averaging is performed over the coarse- and fine-grid template banks (i.e. the respective Wigner-Seitz cells).

The probability distribution of signal mismatches in a given template bank constructed with a certain maximal mismatch m depends on the structure and dimensionality of the template bank. The corresponding average mismatch can be expressed as $\langle \mu \rangle = \xi m$, where $\xi \in (0, 1)$ is a characteristic geometric factor of the template bank. Such mismatch distributions were studied quantitatively, for example in [55]. For hyper-cubic lattices, the geometric relation is well known to be exactly $\langle \mu \rangle = m/3$, which was used in previous optimization studies [64, 88]. For the more efficient A_n^* -lattices this geometric factor is approximately $\xi \approx 0.5 - 0.6$ for low dimensions $n \lesssim 6$. Here we allow for general geometric factors ξ , but for simplicity we assume it to be identical for the fine- and coarse-grid template banks, and so Eq. (4.22) can be written as

$$\langle \hat{\mu} \rangle = \xi(\hat{m} + \tilde{m}), \quad \text{with} \quad \xi \in (0, 1), \quad (4.23)$$

where $\hat{m}$ and $\tilde{m}$ are the maximal mismatch parameters of fine- and coarse-grid template banks, respectively.

Averaging the non-centrality parameter ρ_Σ^2 of Eq. (4.19) over random signal parameters λ_s at fixed signal strength h_{rms} , we can now obtain the expression

$$\langle \rho_\Sigma^2 \rangle = [1 - \xi(\hat{m} + \tilde{m})] \frac{2N_{\text{det}}}{S_n} h_{\text{rms}}^2 T, \quad (4.24)$$

where we (trivially) generalized the result to the case of a network of N_{det} detectors. In this case S_n refers to the harmonic mean over individual-detector PSDs, and h_{rms} is a noise-weighted average over rms-amplitudes from different detectors (e.g. see [44]). The fact that *both* the coarse- and fine-grid mismatches enter this expression has been overlooked in previous studies [64, 88], where only the fine-grid mismatch $\hat{m}$ had been included¹.

4.2.3. Sensitivity estimate

The false-alarm and false-dismissal probabilities for a given threshold Σ_{th} of the StackSlide statistic Σ of Eq. (4.2) are

$$p_{\text{fA}}(\Sigma_{\text{th}}) = \int_{\Sigma_{\text{th}}}^{\infty} \chi_{4N}^2(\Sigma; 0) d\Sigma, \quad (4.25)$$

$$p_{\text{fD}}(\Sigma_{\text{th}}, \rho_\Sigma^2) = \int_{-\infty}^{\Sigma_{\text{th}}} \chi_{4N}^2(\Sigma; \rho_\Sigma^2) d\Sigma, \quad (4.26)$$

where the special case of a coherent $\mathcal{F}$ -statistic search corresponds to $N = 1$.

Sensitivity is often quantified in terms of the weakest (rms-) signal strength h_{th} required to obtain a given detection probability $p_{\text{det}}^* = 1 - p_{\text{fD}}^*$ at a given false-alarm

¹These studies additionally imposed the *ad-hoc* constraint of $\tilde{m} = \hat{m}$ in the computing-cost expressions

probability p_{fA}^* . This requires inverting Eq. (4.25) to obtain the critical threshold $\Sigma_{\text{th}}^* = \Sigma_{\text{th}}(p_{\text{fA}}^*)$, then substituting this into Eq. (4.26) and inverting to find the critical non-centrality parameter

$$\rho_{\Sigma}^{*2} = \rho_{\Sigma}^2(p_{\text{fA}}^*, p_{\text{fD}}^*, N). \quad (4.27)$$

The signal location λ_s is generally unknown, therefore the mismatch $\hat{\mu}(\lambda_s, \hat{\lambda})$ of the closest template $\hat{\lambda}$ and the corresponding mismatched non-centrality parameter $\rho_{\Sigma}^2(\lambda_s, \hat{\lambda})$ of Eq. (4.19) follow a random distribution. In order to estimate the threshold rms signal strength h_{th} , one would have to compute $p_{\text{fD}}(p_{\text{fA}}^*, h_{\text{th}})$ by *averaging* the right-hand side of Eq. (4.26) over the (known) mismatch distribution of $\hat{\mu}$. Furthermore, for statements about physical upper limits and sensitivity of a given search pipeline, it is often required to quantify the sensitivity in terms of the *intrinsic* GW amplitude h_0 , instead of the rms detector strain h_{rms} , which would require further averaging of Eq. (4.26) over the (potentially) unknown sky-position and polarization parameters. This problem has recently been studied in detail in [94].

For our present purpose it will be sufficient to obtain the correct *scaling* of sensitivity with StackSlide parameters $\{N, T, \tilde{m}, \hat{m}\}$, while the absolute sensitivity level is less important. We will therefore employ the usual simplification of this problem, which consists in averaging ρ_{Σ}^2 instead of $p_{\text{fD}}(\rho_{\Sigma}^2)$ over the mismatch distribution of $\hat{\mu}$, so we approximate

$$\begin{aligned} p_{\text{fD}}(p_{\text{fA}}^*, h_{\text{th}}) &= \left\langle p_{\text{fD}}(p_{\text{fA}}^*, \rho_{\Sigma}^2) \Big|_{h_{\text{th}}} \right\rangle_{\lambda_s} \\ &\approx p_{\text{fD}}\left(p_{\text{fA}}^*, \langle \rho_{\Sigma}^2 \rangle_{\lambda_s}\right). \end{aligned} \quad (4.28)$$

The results of [94] indicate that this indeed approximately preserves the *scaling* of sensitivity as a function of StackSlide parameters.

We can now use Eq. (4.24) to translate the critical non-centrality parameter ρ_{Σ}^{*2} of Eq. (4.27) into a threshold rms signal-strength h_{th} , namely

$$h_{\text{th}}^{-2} = \frac{2N_{\text{det}}}{\rho_{\Sigma}^{*2}} [1 - \xi(\hat{m} + \tilde{m})] \frac{T}{S_n}. \quad (4.29)$$

Following the Neyman-Pearson criterion we want to maximize detection probability $p_{\text{det}}^* = 1 - p_{\text{fD}}^*$ at fixed false-alarm probability p_{fA}^* and at fixed signal strength h_{rms} . Equivalently² we can fix the false-alarm and false-dismissal probabilities and *minimize* the required threshold rms signal strength h_{th} , which is the traditional optimization approach used in previous studies [64, 88].

Gauss approximation for large N

One approach (used in [43, 88]) to make further analytical progress consists in assuming a large number of segments, i.e. $N \gg 1$, and invoke the central limit theorem to

²Due the monotonicity of p_{fD} as a function of h_{rms} .

approximate χ_{4N}^2 by a Gaussian distribution

$$P(\Sigma|\rho_\Sigma^2) \stackrel{N \gg 1}{\approx} (2\pi\sigma_\Sigma^2)^{-1/2} \exp\left[-\frac{(\Sigma - \bar{\Sigma})^2}{2\sigma_\Sigma^2}\right], \quad (4.30)$$

with mean and variance of $\chi_{4N}^2(\rho_\Sigma^2)$ given by

$$\begin{aligned} \bar{\Sigma} &= 4N + \rho_\Sigma^2, \\ \sigma_\Sigma^2 &= 2(4N + 2\rho_\Sigma^2). \end{aligned} \quad (4.31)$$

This allows us to analytically integrate Eqs. (4.25), (4.26), which yields

$$p_{\text{fA}}(\Sigma_{\text{th}}) = \frac{1}{2} \operatorname{erfc}\left(\frac{\Sigma_{\text{th}} - 4N}{2\sqrt{4N}}\right), \quad (4.32)$$

$$p_{\text{fD}}(\Sigma_{\text{th}}, \rho_\Sigma^2) = \frac{1}{2} \operatorname{erfc}\left(\frac{\rho_\Sigma^2 - (\Sigma_{\text{th}} - 4N)}{2\sqrt{4N + 2\rho_\Sigma^2}}\right), \quad (4.33)$$

where $\operatorname{erfc}(x) \equiv 1 - \operatorname{erf}(x)$ is the complementary error-function. Substituting Eq. (4.32) into Eq. (4.33), we obtain

$$\beta \equiv \frac{\rho_\Sigma^2 - 2\alpha\sqrt{4N}}{2\sqrt{4N + 2\rho_\Sigma^2}}, \quad (4.34)$$

where we defined

$$\begin{aligned} \beta &\equiv \operatorname{erfc}^{-1}(2p_{\text{fD}}^*) = -\operatorname{erfc}^{-1}(2p_{\text{det}}^*), \\ \alpha &\equiv \operatorname{erfc}^{-1}(2p_{\text{fA}}^*). \end{aligned} \quad (4.35)$$

Solving Eq. (4.34) for the critical non-centrality parameter ρ_Σ^{*2} , we obtain³

$$\rho_\Sigma^{*2}(\alpha, \beta, N) = 2\alpha\sqrt{4N} + 4\beta^2 + 2\beta\sqrt{4N + 4\alpha\sqrt{4N} + 4\beta^2}, \quad (4.36)$$

which we refer to as the ‘‘Gauss approximation’’. Introducing the average per-segment SNR $\rho_{\mathcal{F}}$ as $\rho_{\mathcal{F}}^2 \equiv \langle \rho_\Sigma^2 \rangle / N$, one can consider two interesting limits of the false-dismissal equation (4.34):

- (i) *strong-signal limit* ($\rho_{\mathcal{F}}^2 \gg 1$): the per-segment SNR of the signal is large, and we obtain

$$\rho_\Sigma^* \approx \sqrt{8}\beta, \quad (4.37)$$

which is somewhat pathological, as $\beta \gg 1$ and therefore the detection probability is *extremely* close to $p_{\text{det}} = 1$. Neither false-alarm threshold nor the number of segments N matter for detectability⁴ in this case.

³The second solution has $\beta < 0$, corresponding to $p_{\text{fD}} > 0.5$.

⁴This has been noted previously for radio observations[95]

- (ii) *weak-signal limit* ($\rho_{\mathcal{F}}^2 \ll 1$): the per-segment SNR of the signal is small, and using $N \gg 1$ we find

$$\rho_{\Sigma}^{*2} \approx 2\sqrt{4N}(\alpha + \beta), \quad (4.38)$$

which we refer to as the “weak-signal Gauss approximation” (WSG), which was first used in [43] to estimate the sensitivity of the Hough method. This approach results in the “classic” semicoherent sensitivity scaling as a function of N , namely

$$h_{\text{th}, \text{WSG}}^{-2} = \frac{N_{\text{det}}}{2S_n} \frac{[1 - \xi(\hat{m} + \tilde{m})]}{\alpha + \beta} \frac{T}{\sqrt{N}}. \quad (4.39)$$

In practice we find that the WSG approximation is often not well satisfied, and the deviations of the N -scaling in Eq. (4.38) from the exact form of Eq. (4.27) can lead to dramatically different optimal solutions. Already the Gauss approximation of Eq. (4.36) is not well satisfied for small false-alarm probabilities $p_{\text{fA}} \ll 1$ and segment numbers in the range $N \lesssim \mathcal{O}(1000)$, as can be seen in Fig. 4.2. A more reliable approximation was recently introduced in [94], namely using the Gaussian distribution only for the false-dismissal equation [(4.26)], while keeping the central χ^2 -distribution for the false-alarm equation [(4.25)]. For the present chapter this approach would not be well-suited, however, as we need the sensitivity equation in the form of a power-law in T and N , similarly to Eq. (4.39).

Local power-law approximation for ρ_{Σ}^*

We can incorporate the exact N -scaling of the critical non-centrality parameter ρ_{Σ}^{*2} of Eq. (4.27) by *locally* expressing it as a power-law in the form

$$\rho_{\Sigma}^{*2}(p_{\text{fA}}^*, p_{\text{fD}}^*, N) = r_0 N^{1/(2w)}, \quad (4.40)$$

where $w(p_{\text{fA}}^*, p_{\text{fD}}^*, N_0)$ is a parameter quantifying the relative deviation of the exact N -scaling from the WSG limit of Eq. (4.38), where $w = 1$. The power-law coefficients can be computed as

$$w = \left(2 \frac{\partial \log \rho_{\Sigma}^{*2}}{\partial \log N} \right)^{-1}, \quad r_0 = \rho_{\Sigma}^{*2} N_0^{-1/(2w)}, \quad (4.41)$$

evaluated at a point $\{p_{\text{fA}}^*, p_{\text{fD}}^*, N_0\}$.

The function $w(N)$ is shown in Fig. 4.2, for a reference false-dismissal probability of $p_{\text{fD}}^* = 0.1$ and different choices of false-alarm probability p_{fA}^* , both for the exact solution Eq. (4.27) and for the Gauss approximation of Eq. (4.36). We see that the exact N -scaling w increasingly deviates from the WSG approximation ($w = 1$) at *lower* false-alarm probabilities and at smaller N . The Gauss approximation tends to agree better with the exact scaling at larger N (as expected), and at *higher* false-alarm probabilities.

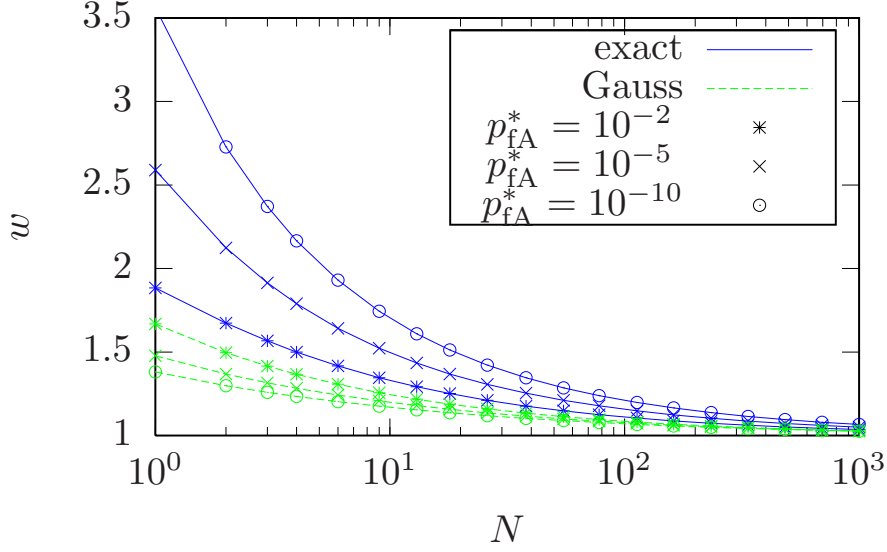


Figure 4.2.: N -scaling coefficient w defined in Eq. (4.40) as a function of N , for false-dismissal probability $p_{\text{fD}}^* = 0.1$, and different false-alarm probabilities $p_{\text{fA}}^* \in [10^{-10}, 10^{-5}, 10^{-2}]$. Solid lines show the scaling obtained from the exact solution Eq. (4.27), while dashed lines refer to the Gauss approximation of Eq. (4.36). The WSG approximation corresponds to $w = 1$.

Using the power-law approximation of Eq. (4.40), we can now express the threshold signal strength of Eq. (4.29) as

$$\begin{aligned} L_0(N, T, \tilde{m}, \hat{m}) &\equiv \frac{r_0 S_n}{2N_{\text{det}}} h_{\text{th}}^{-2} \\ &= [1 - \xi(\tilde{m} + \hat{m})] T N^{-1/(2w)}, \end{aligned} \quad (4.42)$$

which defines the objective function L_0 that we want to maximize as a function of the StackSlide parameters.

We see that, without further constraints the optimal solution would simply be $m \rightarrow 0$, $N \rightarrow 1$ and $T \rightarrow T_{\text{max}}$, i.e. a fully coherent search over all the available data T_{max} with an infinitely fine template bank. This would obviously require infinite computing power, and we therefore need to extend the optimization problem by a computing-cost constraint.

4.2.4. Template counting

For both the coarse⁵ and the fine grid, the respective number of templates⁶ $\{\tilde{\mathcal{N}}, \hat{\mathcal{N}}\}$ covering the parameter space $\mathbb{P}$ is given [54, 55] by the general expression

$$\mathcal{N} = \theta_n m^{-n/2} \mathcal{V}_n, \quad \text{with} \quad \mathcal{V}_n \equiv \int_{\mathbb{T}_n} d^n \lambda \sqrt{\det g}, \quad (4.43)$$

where m is the maximal-mismatch parameter, $\det g$ is the determinant of the corresponding parameter-space metric g_{ij} , and $\mathcal{V}_n$ denotes the metric volume of the n -dimensional space $\mathbb{T}_n \subseteq \mathbb{P}$ spanned by the template-bank. The normalized thickness θ_n corresponding to the number of templates per unit volume [54], depends on the geometric structure of the covering, for example $\theta_{\mathbb{Z}_n} = n^{n/2} 2^{-n}$ for a hyper-cubic lattice $\mathbb{Z}_n$.

An important subtlety in Eq. (4.43) is the dimensionality n of the *template-bank space* $\mathbb{T}_n$, which can be smaller than the dimensionality of the parameter space $\mathbb{P}$, as previously discussed in [64, 88]. The template-bank dimensionality n is generally a (piece-wise constant) function of the StackSlide parameters $\{N, \Delta T, m\}$, which determine the metric resolution. The extent of $\mathbb{P}$ along certain directions can be “thin” compared to the metric resolution and would require only a single template along this direction, effectively not contributing to the template-bank dimensionality. For different StackSlide parameters, however, the resolution might be sufficient to require more than one template along this direction, adding to the template-bank dimensionality n .

Following [53, 64, 88], the correct dimensionality for given StackSlide parameters can be determined by the condition that n should *maximize* the number of templates $\mathcal{N}_n$ computed via Eq. (4.43), i.e.,

$$\tilde{\mathcal{N}}_{\tilde{n}} = \max_n \tilde{\mathcal{N}}_n, \quad \text{and} \quad \hat{\mathcal{N}}_{\hat{n}} = \max_n \hat{\mathcal{N}}_n. \quad (4.44)$$

This can be understood as follows: if $\mathcal{N}$ decreases when adding a template-bank dimension, then the corresponding parameter-space extent is thinner than the metric resolution and therefore adds “fractional” templates. On the other hand, if $\mathcal{N}$ decreases by removing a dimension, then its extent is thicker than the metric resolution and requires more than one template to cover it.

An interesting alternative formulation can be obtained by expressing Eq. (4.44) as the condition $\mathcal{N}_n / \mathcal{N}_{n-1} > 1$ for including an additional dimension n . For constant metrics and simple parameter-space shapes, i.e. $\mathcal{V}_n = \int d^n \lambda \sqrt{g} = \sqrt{g} \Delta \lambda_1 \times \Delta \lambda_2 \dots \times \Delta \lambda_n$, this can be shown to be equivalent to

$$\frac{\theta_n}{\theta_{n-1}} \frac{\Delta \lambda_n}{d \lambda_n} > 1, \quad (4.45)$$

⁵We assume a roughly constant number of coarse-grid templates $\tilde{\mathcal{N}}$ across all segments.

⁶The *templates* in this formulation are not to be confused with the “patches” used in BC [64] and CGK [88]. A “patch” in the BC/CGK framework corresponds to a *line* of templates along the frequency axis.

where $d\lambda_n \equiv \sqrt{m g^{nn}}$ is the metric *template extent* along dimension n , in terms of the diagonal element g^{nn} of the inverse metric g^{ij} . This shows that Eq. (4.44) boils down to (apart from the lattice-thickness ratio) the requirement that the parameter-space extent $\Delta\lambda_n$ along a given dimension n must exceed the corresponding metric template resolution $d\lambda_n$.

The coherent (coarse-grid) metric volume $\tilde{\mathcal{V}}_{\tilde{n}}$ is typically a steep function of the coherence time ΔT , and can often be well approximated (over some range of ΔT) by a power law, namely $\tilde{\mathcal{V}}_{\tilde{n}}(\Delta T) \propto \Delta T^{\tilde{q}}$. We can therefore write Eq. (4.43) for $\tilde{\mathcal{N}}$ in the power-law form

$$\tilde{\mathcal{N}}_{\tilde{n}}(\tilde{m}, \Delta T) = \tilde{k} \tilde{m}^{-\tilde{n}/2} \Delta T^{\tilde{q}}, \quad (4.46)$$

where $\tilde{k} = \theta_{\tilde{n}} \tilde{\mathcal{V}}_{\tilde{n}}(\Delta T_0) \Delta T_0^{-\tilde{q}}$ for some choice of segment length ΔT_0 .

The semicoherent (fine-grid) metric volume $\hat{\mathcal{V}}_{\hat{n}}$ generally depends on both ΔT and N and can typically [64, 56, 93] be factored in the form

$$\hat{\mathcal{V}}_{\hat{n}}(N, \Delta T) = \gamma_{\hat{n}}(N) \tilde{\mathcal{V}}_{\hat{n}}(\Delta T), \quad (4.47)$$

in terms of the *refinement factor* $\gamma_{\hat{n}}(N) \geq 1$ and the coherent-metric volume $\tilde{\mathcal{V}}_{\hat{n}}(\Delta T)$ of the fine-grid template space. Typically $\gamma(N)$ can be well approximated (over some range of N) by a power law, namely $\gamma(N) \propto N^{\hat{p}}$. We can therefore write Eq. (4.43) for $\hat{\mathcal{N}}$ in the power-law form

$$\hat{\mathcal{N}}_{\hat{n}}(\hat{m}, \Delta T, N) = \hat{k} \hat{m}^{-\hat{n}/2} \Delta T^{\hat{q}} N^{\hat{p}}, \quad (4.48)$$

where $\hat{k} = \theta_{\hat{n}} \hat{\mathcal{V}}_{\hat{n}}(\Delta T_0, N_0) \Delta T_0^{-\hat{q}} N_0^{-\hat{p}}$ for some choice of parameters $\{\Delta T_0, N_0\}$.

4.2.5. Computing-cost model

The total computing cost C_{tot} of the interpolating StackSlide statistic has two main contributions, namely

$$C_{\text{tot}}(\tilde{m}, \hat{m}, N, \Delta T) = \tilde{C} + \hat{C}, \quad (4.49)$$

where $\tilde{C}(\tilde{m}, N, \Delta T)$ is the computing cost of the $\mathcal{F}$ -statistic over the coarse grid of $\tilde{\mathcal{N}}_{\tilde{n}}$ templates for each of the N segments, and $\hat{C}(\hat{m}, N, \Delta T)$ is the cost of incoherently summing these $\mathcal{F}$ -values across all segments on a fine grid of $\hat{\mathcal{N}}_{\hat{n}}$ templates. Note that we neglect all other costs such as data-IO etc, which for any computationally limited search will typically be much smaller than C_{tot} .

Computing cost $\tilde{C}$ of the coherent step

The computing cost of the coherent step is

$$\tilde{C}(\tilde{m}, N, \Delta T) = N \tilde{\mathcal{N}}_{\tilde{n}}(\tilde{m}, \Delta T) N_{\text{det}} \tilde{c}_1(\Delta T), \quad (4.50)$$

where $\tilde{c}_1(\Delta T)$ is the $\mathcal{F}$ -statistic computing cost of a single template for a single segment and a single detector. Here we used the fact that to first order [41] the number of detectors has no effect on the number of templates $\tilde{\mathcal{N}}$.

As discussed previously in [88], there are two fundamentally different implementations of the $\mathcal{F}$ -statistic calculation currently in use: a direct *SFT-method* [96], and a (generally far more efficient) *FFT-method* based on barycentric resampling [91, 97].

- (i) The *SFT-method* consists in interpolating frequency bins of short Fourier transforms (“SFTs”) of length T_{SFT} , using approximations described in [96, 44]. The resulting per-template cost $\tilde{c}_1(\Delta T)$ is directly proportional to the segment length ΔT :

$$\tilde{c}_1^{\text{SFT}}(\Delta T) = \tilde{c}_0^{\text{SFT}} \frac{\Delta T}{T_{\text{SFT}}}, \quad (4.51)$$

where $\tilde{c}_0^{\text{SFT}}$ is an implementation- and hardware-dependent fundamental computing cost.

- (ii) In the *FFT-method* the cost of searching a frequency band Δf using an (up-sampled by u) FFT frequency-resolution of $u/\Delta T$ is proportional to $\mathcal{N}_f \log 2\mathcal{N}_f$, where $\mathcal{N}_f = u\Delta f \Delta T$ is the number of frequency bins. We can therefore express the per-template $\mathcal{F}$ -statistic cost $\tilde{c}_1(\Delta T)$ as

$$\tilde{c}_1^{\text{FFT}}(\Delta T) = \tilde{c}_0^{\text{FFT}} \log(2u\Delta f \Delta T), \quad (4.52)$$

where $\tilde{c}_0^{\text{FFT}}$ is an implementation- and hardware-dependent fundamental computing cost.

Using the power-law model of Eq. (4.46) for $\tilde{\mathcal{N}}$, we can write the coherent computing cost in the form

$$\tilde{C}(\tilde{m}, N, \Delta T) = \tilde{\kappa} \tilde{m}^{-\tilde{\eta}/2} N^{\tilde{\eta}} \Delta T^{\tilde{\delta}}, \quad (4.53)$$

where

$$\tilde{\eta} = 1, \quad \tilde{\delta} = \tilde{q} + \Delta\tilde{\delta}, \quad (4.54)$$

and where $\Delta\tilde{\delta}$ is either

$$\Delta\tilde{\delta} = \begin{cases} \Delta\tilde{\delta}_{\text{SFT}} \equiv 1, \\ \Delta\tilde{\delta}_{\text{FFT}} \equiv [\log(2u\Delta f \Delta T_0)]^{-1}, \end{cases} \quad (4.55)$$

depending on whether the $\mathcal{F}$ -statistic is computed using the *SFT-* or *FFT-method*, respectively. The expression for $\Delta\tilde{\delta}_{\text{FFT}}$ can be obtained via Eq. (4.62) and depends (albeit weakly) on the reference segment length ΔT_0 . The corresponding proportionality factors $\tilde{\kappa}$ are found as

$$\begin{aligned} \tilde{\kappa}_{\text{SFT}} &= \theta_{\tilde{n}} N_{\text{det}} \frac{\tilde{c}_0^{\text{SFT}}}{T_{\text{SFT}}} \frac{\tilde{\mathcal{V}}_{\tilde{n}}(\Delta T_0)}{\Delta T_0^{\tilde{q}}}, \\ \tilde{\kappa}_{\text{FFT}} &= \theta_{\tilde{n}} N_{\text{det}} \frac{\tilde{c}_0^{\text{FFT}}}{\Delta\tilde{\delta}_{\text{FFT}}} \frac{\tilde{\mathcal{V}}_{\tilde{n}}(\Delta T_0)}{\Delta T_0^{\tilde{\delta}}}. \end{aligned} \quad (4.56)$$

Computing cost $\hat{C}$ of the incoherent step

The computing cost of the StackSlide step is

$$\hat{C}(\hat{m}, N, \Delta T) = N \hat{\mathcal{N}}_{\hat{n}}(\hat{m}, \Delta T, N) \hat{c}_0, \quad (4.57)$$

where $\hat{c}_0$ is the implementation- and hardware-dependent fundamental cost of adding one value of $2\mathcal{F}_k$ for one fine-grid point $\hat{\lambda}$ in Eq. (4.2), including the cost of mapping the fine-grid point to its closest coarse-grid template $\hat{\lambda}_k(\hat{\lambda})$ using the coherent metric of the segment k , namely $\tilde{g}_k$. The incoherent step operates on coherent multi-detector $\mathcal{F}$ -statistic values, and therefore does not depend on the number of detectors N_{det} .

Using the power-law model of Eq. (4.48) for $\hat{\mathcal{N}}$, we can write the incoherent computing cost as

$$\hat{C}(\hat{m}, N, \Delta T) = \hat{\kappa} \hat{m}^{-\hat{n}/2} N^{\hat{\eta}} \Delta T^{\hat{\delta}}, \quad (4.58)$$

where

$$\hat{\eta} = \hat{p} + 1, \quad \hat{\delta} = \hat{q}, \quad (4.59)$$

and the proportionality factor

$$\hat{\kappa} = \theta_{\hat{n}} \hat{c}_0 \frac{\hat{\mathcal{V}}_{\hat{n}}(N_0, \Delta T_0)}{N_0^{\hat{p}} \Delta T_0^{\hat{q}}}, \quad (4.60)$$

for given reference values $\{N_0, \Delta T_0\}$.

General power-law computing-cost model

Combining Eqs. (4.53) and (4.58) we arrive at the following power-law model for the total computing cost, defined in Eq. (4.49), namely

$$C_{\text{tot}} = \tilde{\kappa} \tilde{m}^{-\tilde{n}/2} N^{\tilde{\eta}} \Delta T^{\tilde{\delta}} + \hat{\kappa} \hat{m}^{-\hat{n}/2} N^{\hat{\eta}} \Delta T^{\hat{\delta}}. \quad (4.61)$$

If a given computing-cost function does not follow this model, we can always produce a local fit to Eq. (4.61), which should be valid over some range of parameters $\{\Delta T, N\}$, namely

$$\delta \equiv \frac{\partial \log C}{\partial \log \Delta T}, \quad \eta \equiv \frac{\partial \log C}{\partial \log N}, \quad (4.62)$$

$$\kappa \equiv \frac{C(m_0, N_0, \Delta T_0)}{m_0^{-n/2} N_0^{\eta} \Delta T_0^{\delta}}, \quad (4.63)$$

for reference values $\{m_0, N_0, \Delta T_0\}$. Note that δ generally depends only on ΔT_0 , while η depends only on N_0 , due to the way these dependencies typically factor (cf. Sec. 4.2.5). The mismatch dependency $\propto m^{-n/2}$ is exact according to Eq. (4.43), but a given computing-cost function might still deviate from this behaviour (e.g. the BC/CGK

computing-cost function discussed in Sec. 4.4.3). In this case one can extend the power-law fit by extracting the “mismatch-dimension” n via

$$n \equiv -2 \frac{\partial \log C}{\partial \log m} . \quad (4.64)$$

It will be more convenient in the following to work in terms of $\{N, T\}$ instead of $\{N, \Delta T\}$, where $T = N \Delta T$ is the total time span of data used. Changing variables, we obtain the computing-cost model in the form

$$C_{\text{tot}} = \tilde{\kappa} \tilde{m}^{-\tilde{n}/2} N^{-\tilde{\varepsilon}} T^{\tilde{\delta}} + \hat{\kappa} \hat{m}^{-\hat{n}/2} N^{-\hat{\varepsilon}} T^{\hat{\delta}} , \quad (4.65)$$

where we defined

$$\varepsilon \equiv \delta - \eta , \quad (4.66)$$

generally satisfying $\varepsilon > 0$ for all realistic cases considered here. Note that m and N are dimensionless, therefore the respective units of $[C/\kappa]$ are $[T^\delta]$.

Symbol	Description	Relations	Refs
N	Number of segments		Sec. 4.2, Eq. (4.1)
ΔT	Segment duration		Sec. 4.2
T	Total observation time	$T = N \Delta T$	Sec. 4.2
$\tilde{Q}$	a quantity Q referring to the <i>coherent</i> step		Sec. 4.2
$\hat{Q}$	a quantity Q referring to the <i>incoherent</i> step		Sec. 4.2
n	number of template-bank dimensions		Eq. (4.44)
m	maximal template-bank mismatch parameter		Eqs. (4.8), (4.15)
ξ	average mismatch factor $\in [0, 1]$	$\langle \mu \rangle = \xi m$	Eq. (4.23)
w	L_0 sensitivity scaling with N	$L_0 _{m,T} \propto N^{-1/(2w)}$	Eq. (4.40)
ϖ	Lagrange multiplier for computing-cost constraint	$L = L_0 + \varpi(C_{\text{tot}} - C_0)$	Eq. (4.67)
κ	computing-cost prefactor		Eqs. (4.53), (4.58)
δ	computing-cost T - or ΔT - exponent at fixed N	$C _N \propto \Delta T^\delta \propto T^\delta$	Eqs. (4.53), (4.58)
η	computing-cost N -exponent at fixed ΔT	$C _{\Delta T} \propto N^\eta$	Eqs. (4.53), (4.58)
$-\varepsilon$	computing-cost N -exponent at fixed T	$C _T \propto N^{-\varepsilon}$	Eq. (4.65)

Table 4.1.: Overview of symbols and notation used in the formulation of the optimization problem.

4.3. Maximizing sensitivity at fixed computing cost

We want to maximize the objective function $L_0 \propto h_{\text{th}}^{-2}$ defined in Eq. (4.42) under the constraint of fixed computing cost, $C_{\text{tot}} = C_0$. We therefore need to find the stationary points of the Lagrange function

$$L(N, T, \tilde{m}, \hat{m}, \varpi) = L_0 + \varpi [\tilde{C} + \hat{C} - C_0] , \quad (4.67)$$

where stationarity with respect to the Lagrange multiplier, i.e. $\partial_\varpi L = 0$, returns the computing-cost constraint $\tilde{C} + \hat{C} = C_0$.

Table 4.1 provides a “dictionary” summarizing the notation used here and in the previous section to formulate the optimization problem.

Before embarking on the full optimization problem, it is instructive to consider two special cases, namely (i) a fully coherent search, and (ii) searches where the computing cost is dominated by one contribution, either coherent $\tilde{C}$ or incoherent $\hat{C}$.

4.3.1. Special case (i): Fully coherent search

The fully coherent search is a special case of Eq. (4.67) with the additional constraint $N = 1$, and therefore $\Delta T = T$, $\hat{m} = 0$, and $\hat{C} = 0$. This leaves us with the reduced Lagrangian

$$L(T, \tilde{m}, \varpi) = (1 - \xi \tilde{m}) T + \varpi [\tilde{\kappa} \tilde{m}^{-\tilde{n}/2} T^{\tilde{\delta}} - C_0]. \quad (4.68)$$

Requiring stationarity with respect to $\{T, \tilde{m}, \varpi\}$ results in the optimal StackSlide parameters

$$\xi \tilde{m}_{\text{opt}} = \left(1 + \frac{2\tilde{\delta}}{\tilde{n}}\right)^{-1}, \quad (4.69)$$

$$T_{\text{opt}} = \left(\frac{C_0}{\tilde{\kappa}}\right)^{1/\tilde{\delta}} \tilde{m}_{\text{opt}}^{\tilde{n}/(2\tilde{\delta})}, \quad (4.70)$$

Interestingly the optimal mismatch $\tilde{m}_{\text{opt}}$ is independent of both the computing-cost constraint C_0 and the observation time T . The scaling of the resulting threshold signal strength h_{th} with computing cost C_0 is therefore

$$h_{\text{th}}^{-1} \propto C_0^{1/(2\tilde{\delta})}. \quad (4.71)$$

In practical applications we often find $\tilde{\delta} \approx 3 - 7$, and so T_{opt} and h_{th}^{-1} will increase very slowly with increasing computing cost C_0 . This indicates that a brute-force approach of throwing more computing power at a fully coherent search will typically yield meagre returns in sensitivity.

4.3.2. Special case (ii): Computing cost dominated by one contribution

If either the coherent $\tilde{C}$ or incoherent $\hat{C}$ contribution dominates the total computing cost [(4.65)], we can write

$$C_{\text{tot}} \approx \kappa m^{-n/2} N^{-\varepsilon} T^{\delta}, \quad (4.72)$$

where all StackSlide parameters now refer to dominant contribution only.

We assume that the negligible computing-cost contribution implies that we can also neglect the corresponding mismatch: if the respective step is cheap, one can easily increase sensitivity by reducing the corresponding mismatch until it is negligible, i.e. we assume $\langle \mu_{\text{sc}} \rangle \approx \xi m$. This qualitative argument will be confirmed by the general solution in the next section. We can therefore write the objective function Eq. (4.42) as

$$L_0(N, T, m) \approx (1 - \xi m) N^{-1/(2w)} T. \quad (4.73)$$

Using Eq. (4.72) we can obtain

$$N(C_0, m, T) = (C_0/\kappa)^{-1/\varepsilon} m^{-n/(2\varepsilon)} T^{\delta/\varepsilon}, \quad (4.74)$$

$$\Delta T(C_0, m, T) = (C_0/\kappa)^{1/\varepsilon} m^{n/(2\varepsilon)} T^{-\eta/\varepsilon}, \quad (4.75)$$

which shows that increasing T at fixed C_0 results in more and shorter segments, while increasing C_0 at fixed T results in fewer and longer segments (assuming $\varepsilon > 0$). Substituting this into Eq. (4.73) yields the threshold signal strength

$$h_{\text{th}}^{-2} \propto (C_0/\kappa)^{1/(2w\varepsilon)} [(1 - \xi m) m^{n/(4w\varepsilon)}] T^{a/(2w\varepsilon)}, \quad (4.76)$$

where we introduced the parameter

$$a \equiv 2w\varepsilon - \delta, \quad (4.77)$$

which will be of critical importance in determining the character of the optimal solution.

The objective function $L_0 \propto h_{\text{th}}^{-2}$ can be easily maximized over mismatch m , resulting in

$$\xi m_{\text{opt}}^{(0)} = \left[1 + \frac{4w\varepsilon}{n} \right]^{-1}, \quad (4.78)$$

which is independent of both C_0 and T . This solution differs from Eq. (4.69) of the fully coherent case, even when the coherent cost dominates (where $\tilde{\varepsilon} = \tilde{\delta} - 1$).

We see in Eq. (4.76) that there is *no extremum* of h_{th} (at least in regions of approximately constant power-law exponents). Given that $w \geq 1$ and generally $\varepsilon > 0$, we can distinguish two different regimes depending on the sign of critical scaling exponent a defined in Eq. (4.77):

$a > 0$: sensitivity improves (i.e. h_{th}^{-1} increases) with *increasing* T (at fixed C_0). Therefore sensitivity is only limited by the total amount of data T_{max} available.

$a < 0$: sensitivity improves (i.e. h_{th}^{-1} decreases) with *decreasing* T , so one should use less data (until the assumptions change).

In practice these extreme conclusions will be modified, as the power-law exponents will vary (slowly) as functions of N and T , and the assumption of a dominating computing-cost contribution might also no longer be satisfied. The marginal case $a = 0$ marks a possible sensitivity maximum, namely if increasing T results in $a < 0$ and decreasing T leads to $a > 0$.

We can obtain a useful qualitative picture of the full optimization problem by considering the two extreme cases of dominating computing contribution $\tilde{C}$ or $\hat{C}$:

- if $\tilde{C} \gg \hat{C}$: we always have $\tilde{a} > 0$ (for all cases of interest $\tilde{\eta} = 1$, $\tilde{\delta} > 2$ and $w \geq 1$). Therefore sensitivity improves with increasing T . As seen in Sec. 4.3.3 this shifts computing cost to the incoherent contribution. Eventually one either uses all the data T_{max} or the coherent cost no longer dominates.

- if $\widehat{C} \gg \widetilde{C}$: the incoherent parameter $\widehat{a}$ can have any sign. If $\widehat{a} > 0$ one would *increase* T until all the data $T_{\max}$ is used (or we reach $\widehat{a} = 0$). If $\widehat{a} < 0$ one would *decrease* T until the incoherent cost no longer dominates.

These limiting cases show that the *type* of optimal solution will be determined solely by the *incoherent* critical exponent $\widehat{a} = 2w\widehat{\varepsilon} - \widehat{\delta}$, namely

$$T_{\text{opt}}^{(0)} = \begin{cases} \text{finite} & \text{if } \widehat{a} \leq 0, \\ \infty & \text{otherwise,} \end{cases} \quad (4.79)$$

which we refer to as the *bounded* and the *unbounded* regime, respectively.

4.3.3. General optimality conditions

We now return to the full optimization problem of Eq. (4.67), namely

$$L = L_0 + \varpi [\widetilde{C} + \widehat{C} - C_0], \quad (4.80)$$

where

$$L_0 = [1 - \xi(\widetilde{m} + \widehat{m})] T N^{-1/(2w)}, \quad (4.81)$$

$$\widetilde{C} = \widetilde{\kappa} \widetilde{m}^{-\widetilde{n}/2} N^{-\widetilde{\varepsilon}} T^{\widetilde{\delta}}, \quad (4.82)$$

$$\widehat{C} = \widehat{\kappa} \widehat{m}^{-\widehat{n}/2} N^{-\widehat{\varepsilon}} T^{\widehat{\delta}}. \quad (4.83)$$

It will be useful introduce the computing-cost *ratio*

$$\varkappa \equiv \widetilde{C}/\widehat{C}, \quad (4.84)$$

and express the respective contributions as

$$\widetilde{C} = \frac{C_0}{1 + \varkappa^{-1}}, \quad \widehat{C} = \frac{C_0}{1 + \varkappa}. \quad (4.85)$$

Using Eqs. (4.82) and (4.83) to solve for T and N , respectively, we obtain

$$N^D = \frac{(C_0/\widehat{\kappa})^{\widetilde{\delta}} [\widetilde{m}^{-\widetilde{n}/2} (1 + \varkappa^{-1})]^{\widetilde{\delta}}}{(C_0/\widetilde{\kappa})^{\widehat{\delta}} [\widehat{m}^{-\widehat{n}/2} (1 + \varkappa)]^{\widehat{\delta}}}, \quad (4.86)$$

$$T^D = \frac{(C_0/\widehat{\kappa})^{\widetilde{\varepsilon}} [\widetilde{m}^{-\widetilde{n}/2} (1 + \varkappa^{-1})]^{\widetilde{\varepsilon}}}{(C_0/\widetilde{\kappa})^{\widehat{\varepsilon}} [\widehat{m}^{-\widehat{n}/2} (1 + \varkappa)]^{\widehat{\varepsilon}}}, \quad (4.87)$$

where D is the determinant of the matrix $[\widetilde{\delta}, \widetilde{\eta}; \widehat{\delta}, \widehat{\eta}]$, which for all cases of practical interest seems to be positive definite, namely

$$D \equiv \widetilde{\delta}\widehat{\eta} - \widehat{\delta}\widetilde{\eta} > 0. \quad (4.88)$$

The segment length $\Delta T = T/N$ can similarly be obtained as

$$\Delta T^D = \frac{(C_0/\widetilde{\kappa})^{\widehat{\eta}} [\widehat{m}^{-\widehat{n}/2} (1 + \varkappa)]^{\widehat{\eta}}}{(C_0/\widehat{\kappa})^{\widetilde{\eta}} [\widetilde{m}^{-\widetilde{n}/2} (1 + \varkappa^{-1})]^{\widetilde{\eta}}}. \quad (4.89)$$

Stationarity with respect to mismatches $\{\tilde{m}, \hat{m}\}$

Requiring stationarity with respect to the mismatches, i.e. $\partial_{\tilde{m}}L = \partial_{\hat{m}}L = 0$, yields

$$\begin{aligned}\varpi\tilde{C} &= -2\xi \frac{\tilde{m}_{\text{opt}}}{\tilde{n}} TN^{-1/(2w)}, \\ \varpi\hat{C} &= -2\xi \frac{\hat{m}_{\text{opt}}}{\hat{n}} TN^{-1/(2w)},\end{aligned}\tag{4.90}$$

which results in the remarkable relation

$$\frac{\tilde{m}_{\text{opt}}/\tilde{n}}{\hat{m}_{\text{opt}}/\hat{n}} = \varkappa.\tag{4.91}$$

The ratio of optimal mismatch per dimension is simply given by the computing-cost ratio $\varkappa$. This result confirms an assumption made in Sec. 4.3.2 about the optimal solution, namely that a negligible computing-cost contributions also implies that one can neglect the corresponding mismatch.

Stationarity with respect to number of segments N

Requiring stationarity with respect to N (treated as continuous), i.e. $\partial_N L = 0$ yields

$$L_0 + 2w\varpi \left[\tilde{\varepsilon}\tilde{C} + \hat{\varepsilon}\hat{C} \right] = 0,\tag{4.92}$$

and substituting Eqs. (4.90) and (4.81), we obtain

$$\frac{\tilde{m}_{\text{opt}}}{\tilde{m}_{\text{opt}}^{(0)}} + \frac{\hat{m}_{\text{opt}}}{\hat{m}_{\text{opt}}^{(0)}} = 1,\tag{4.93}$$

where we used the asymptotic optimal mismatches $m_{\text{opt}}^{(0)}$ defined in Eq. (4.78) for the two limiting cases of dominating coherent or incoherent computing-cost, respectively. Equation (4.93) can be interpreted as defining a two-dimensional ellipse in $\sqrt{m}$ with semi-major axes $\sqrt{m_{\text{opt}}^{(0)}}$. Combining this with Eq. (4.91) we obtain the optimal mismatches

$$\begin{aligned}\frac{\tilde{n}}{\tilde{m}_{\text{opt}}} &= \frac{\tilde{n}}{\tilde{m}_{\text{opt}}^{(0)}} + \frac{\hat{n}}{\hat{m}_{\text{opt}}^{(0)}} \varkappa^{-1}, \\ \frac{\hat{n}}{\hat{m}_{\text{opt}}} &= \frac{\hat{n}}{\hat{m}_{\text{opt}}^{(0)}} + \frac{\tilde{n}}{\tilde{m}_{\text{opt}}^{(0)}} \varkappa,\end{aligned}\tag{4.94}$$

which reduces to the limiting cases of Eq. (4.78) when either computing cost dominates, i.e. when $\varkappa \ll 1$ or $\varkappa \gg 1$. We can express the optimal mismatch prefactor in Eq. (4.81) as

$$[1 - \xi(\tilde{m} + \hat{m})]_{\text{opt}} = \left[1 + \frac{1}{4w} \frac{\tilde{n}\varkappa + \hat{n}}{\tilde{\varepsilon}\varkappa + \hat{\varepsilon}} \right]^{-1}.\tag{4.95}$$

The optimal mismatches Eq. (4.94) only depend on the computing-cost ratio $\varkappa$. Substituting into Eq. (4.87) we therefore obtain a relation of the form $T_0 = T(C_0, \varkappa_{\text{opt}})$ for given observation time T_0 , which can (numerically) be solved for $\varkappa_{\text{opt}} = \varkappa(C_0, T_0)$. Similarly, one could specify N_0 and solve Eq. (4.86) for $\varkappa_{\text{opt}} = \varkappa(C_0, N_0)$. In either case the optimal mismatches are obtained from Eq. (4.94) and the optimal number and length of segments from Eqs. (4.86) and (4.87), fully closing the optimal solution at fixed T .

Monotonicity relations with T

It is interesting to consider the behaviour of the optimal “fixed- T ” solution of the previous section as a function of T . We see in Eq. (4.94) that $\tilde{m}_{\text{opt}}$ is monotonically increasing with $\varkappa$, while $\hat{m}_{\text{opt}}$ is decreasing, i.e.

$$\partial_{\varkappa}\tilde{m}_{\text{opt}} > 0, \quad \text{and} \quad \partial_{\varkappa}\hat{m}_{\text{opt}} < 0. \quad (4.96)$$

We generally assume $D \equiv \tilde{\delta}\hat{\eta} - \hat{\delta}\tilde{\eta} > 0$ and $\varepsilon > 0$ which implies that the right-hand side of Eq. (4.87) is monotonically *decreasing* with $\varkappa$, while the left-hand side is monotonically increasing with T . Therefore $\varkappa$ must be monotonically *decreasing* with T , i.e.

$$\partial_T \varkappa < 0. \quad (4.97)$$

Therefore the optimal solution shifts computing cost from the coherent to the incoherent step with increasing T , which had already been used in Sec. 4.3.2. Combining this with Eq. (4.96) we find

$$\partial_T \tilde{m}_{\text{opt}} < 0, \quad \text{and} \quad \partial_T \hat{m}_{\text{opt}} > 0, \quad (4.98)$$

and using this with Eqs. (4.86) and (4.89), we can further deduce

$$\partial_T N_{\text{opt}} > 0, \quad \text{and} \quad \partial_T \Delta T_{\text{opt}} < 0, \quad (4.99)$$

namely increasing T results in more segments of shorter duration.

Stationarity with respect to observation time T

Requiring stationarity of L with respect to T , i.e. $\partial_T L = 0$, yields the final condition

$$L_0 + \varpi [\tilde{\delta}\tilde{C} + \hat{\delta}\hat{C}] = 0, \quad (4.100)$$

which combined with Eq. (4.92) results in

$$\tilde{a}\tilde{C} + \hat{a}\hat{C} = 0, \quad (4.101)$$

where the critical exponents a are defined in Eq. (4.77). We generally expect $\tilde{a} > 0$, as discussed in Sec. 4.3.2, and therefore the stationarity condition can only have a solution if

$$\hat{a} \equiv \hat{\delta} - 2\hat{\eta} + 2(w-1)\hat{\varepsilon} < 0. \quad (4.102)$$

This conclusion is consistent with the analysis of Sec. 4.3.2: $\hat{a} < 0$ characterizes a *bounded* regime with finite optimal $T_{\text{opt}}^{(0)}$, while $\hat{a} > 0$ characterizes an *unbounded* regime with $T_{\text{opt}}^{(0)} \rightarrow \infty$.

If $T_{\text{opt}}^{(0)}$ exceeds the available data T_{max} , then we simply apply the fixed- T solution of Sec. 4.3.3. Otherwise Eq. (4.101) directly yields

$$\varkappa_{\text{opt}} = -\frac{\hat{a}}{\tilde{a}}, \quad (4.103)$$

closing the optimal solution via Eqs. (4.94), (4.86) and (4.87).

Monotonicity relations with C_0

For a bounded optimal solution with $T_{\text{opt}}^{(0)} \leq T_{\text{max}}$, we see from Eq. (4.103) that $\varkappa_{\text{opt}}$ and $\{\tilde{m}_{\text{opt}}, \hat{m}_{\text{opt}}\}$ are independent of the computing-cost constraint C_0 . Inserting Eqs. (4.86), (4.87) into Eq. (4.42), we can therefore read off the scaling

$$h_{\text{th}}^{-1} \propto C_0^{(\tilde{a}-\hat{a})/(4wD)}, \quad (4.104)$$

which shows that any “reasonable” search should satisfy

$$\tilde{a} > \hat{a}, \quad (4.105)$$

in order for sensitivity to *improve* with increasing C_0 (assuming $D > 0$). Furthermore, from Eqs. (4.86), (4.87) and (4.42) we obtain the monotonicity relations:

$$\begin{aligned} \partial_{C_0} N_{\text{opt}} &\propto \tilde{\delta} - \hat{\delta}, \\ \partial_{C_0} T_{\text{opt}} &\propto \tilde{\varepsilon} - \hat{\varepsilon}, \\ \partial_{C_0} \Delta T_{\text{opt}} &\propto \tilde{\eta} - \hat{\eta}. \end{aligned} \quad (4.106)$$

We expect $\hat{\eta} > \tilde{\eta} = 1$, therefore the optimal segment length ΔT_{opt} will generally increase with C_0 .

The behaviour of the optimal number of segments is less clear-cut: if $\tilde{\delta} < \hat{\delta}$ then N_{opt} *decreases* with C_0 , which can result in a fully coherent search being optimal, despite $T_{\text{opt}}^{(0)} < T_{\text{max}}$. A StackSlide search is therefore not guaranteed to be more sensitive than a fully coherent search at the same computing power, even when computationally limited.

Similarly, T_{opt} can either increase with C_0 (if $\tilde{\varepsilon} > \hat{\varepsilon}$), or decrease: a more expensive and more sensitive search can be using less data.

4.4. Examples of practical application

In order to illustrate the practical application of this analytical framework and its potential gains in sensitivity we consider a few different examples of CW searches.

4.4.1. Directed searches for isolated neutron stars

Directed searches target NSs with *known* sky-position but unknown frequency and frequency derivatives, i.e. $\{f, \dot{f}, \ddot{f}, \dots\}$. The approximate phase metric of this parameter space for isolated NSs is known analytically and constant over the parameter space, e.g. see [Eq. (10) in [62]]. The number of coarse-grid templates scales as

$$\tilde{\mathcal{N}} \propto \Delta T^{n(n+1)/2}, \quad (4.107)$$

while the refinement of the semicoherent metric [Eq. (92) in [56]] scales as

$$\gamma_n \propto N^{n(n-1)/2}. \quad (4.108)$$

The coherent computing-cost exponents Eq. (4.54) are therefore

$$\tilde{\delta} = \tilde{n}(\tilde{n} + 1)/2 + \Delta\tilde{\delta}, \quad \tilde{\eta} = 1, \quad (4.109)$$

where $\Delta\tilde{\delta}$ depends on the $\mathcal{F}$ -statistic implementation as given by Eq. (4.55). The incoherent computing-cost exponents Eq. (4.59) are

$$\hat{\delta} = \hat{n}(\hat{n} + 1)/2, \quad \hat{\eta} = 1 + \hat{n}(\hat{n} - 1)/2, \quad (4.110)$$

which results in $\hat{\varepsilon} = \hat{n} - 1$.

For $\tilde{n} \geq 2$ the condition $\tilde{a} = \tilde{\delta} - 2 + 2(w - 1)\tilde{\varepsilon} > 0$ holds as expected, while the critical boundedness parameter of Eq. (4.102) now reads as

$$\hat{a} = \frac{\hat{n}}{2}(3 - \hat{n}) - 2 + 2(w - 1)(\hat{n} - 1), \quad (4.111)$$

which for the first few values of n evaluates to

$$\begin{aligned} \hat{a}_1 &= -1, \\ \hat{a}_2 &= -3 + 2w, \\ \hat{a}_3 &= -6 + 4w, \\ \hat{a}_4 &= -10 + 6w. \end{aligned} \quad (4.112)$$

In the WSG limit (i.e. $w \rightarrow 1$) this is always $\hat{a} < 0$, and therefore the search falls into the bounded regime. However, in general $w > 1$ and therefore directed StackSlide searches can be either bounded or unbounded.

Directed search for Cassiopeia-A

As a concrete example we consider the directed search for the compact object in Cassiopeia-A (CasA). This search has been performed using LIGO S5 data, and the resulting upper limits have been published in [63]. For the present example we use the search setup as originally proposed in [62], namely a fully coherent $\mathcal{F}$ -statistic search (using the “SFT”

method, i.e. $\Delta\tilde{\delta} = 1$) using data spanning $T = 12$ days, with a maximal template-bank mismatch of $\tilde{m} = 0.2$. The setup assumed two detectors with identical noise floor S_n and a 70% duty cycle, which we can formally incorporate as $N_{\text{det}} = 2 \times 0.7 = 1.4$ in Eqs. (4.50) and (4.29). The parameter space spanned a frequency range of $f \in [100, 300]$ Hz and spindown-ranges corresponding to a spindown-age of $\tau = 300$ y, see [62]. The template-bank dimension for the given StackSlide parameters was determined as $\tilde{n} = 3$, resulting in a power-law scaling of $\tilde{\delta} = 7$ according to Eq. (4.109).

In order to compare sensitivity estimates of different search setups, we use nominal (per-template) false-alarm and false-dismissal probabilities of

$$p_{\text{fA}}^* = 10^{-10}, \quad p_{\text{fD}}^* = 0.1. \quad (4.113)$$

We use a rough estimate of $\xi = 0.5$ (e.g. see [Fig. 8 in [55]]) for the geometric average-mismatch factor of the A_3^* -lattice that was used in this search. Integrating Eqs. (4.25) and (4.26), and solving for $\rho_{\mathcal{F}}^*$ yields $\rho_{\mathcal{F}}^* \approx 8.35$. Substituting this into Eq. (4.29) with $\hat{m} = 0$, $\tilde{m} = 0.2$ yields an estimate for the weakest detectable signal h_{th} of the original $\mathcal{F}$ -statistic search:

$$\left. \frac{h_{\text{th}}}{\sqrt{S_n}} \right|_{\text{ref}} \approx 5.2 \times 10^{-3} \sqrt{\text{Hz}}, \quad (4.114)$$

Timing a current StackSlide code using the SFT-method, one can extract approximate timing parameters

$$\tilde{c}_0^{\text{SFT}} = 7 \times 10^{-8} \text{ s}, \quad \hat{c}_0 = 6 \times 10^{-9} \text{ s}, \quad (4.115)$$

which results in a total computing cost for the original search⁷ of $C_0 \approx 472$ days on a single cluster node. This number is used as the computing-cost constraint C_0 for this example.

First we consider an optimal *coherent* search as described in Sec. 4.3.1, namely using Eq. (4.69) we find

$$\xi \tilde{m}_{\text{opt}} \approx 0.18 \implies \tilde{m}_{\text{opt}} \approx 0.36, \quad (4.116)$$

and using Eq. (4.70) this results in $T_{\text{opt}} = 13.6$ days, which is only about $\sim 13\%$ longer than the original search proposal of [62]. The total improvement in the minimal signal strength h_{th} is less than 2% compared to Eq. (4.114), which shows that the original search proposal was remarkably close to an optimal coherent search.

Next we consider a StackSlide search over the same parameter space using the same computing cost C_0 . Assuming the optimal solution will have segment lengths in the range 1 days $\lesssim \Delta T \lesssim 7$ days, and a total span of $T \lesssim 365$ days, the parameter-space dimensions would be $\tilde{n} = 2$, $\hat{n} = 3$ (see [62]). This results in power-law exponents $\tilde{\delta} = 4$, $\hat{\delta} = 6$, $\hat{\eta} = 4$, and therefore $\tilde{\varepsilon} = 3$, $\hat{\varepsilon} = 2$, and $\tilde{a} = 2 + 6(w - 1)$, $\hat{a} = -2 + 4(w - 1)$. In order to simplify the example we use the WSG approximation, i.e. $w = 1$, which implies

⁷Using the original timing constant $\tilde{c}_{\text{SFT}}^{(0)} = 6 \times 10^{-7} \text{ s}$ of [62], we correctly recover the original estimate of $\tilde{C} \approx 20 \times 200$ days

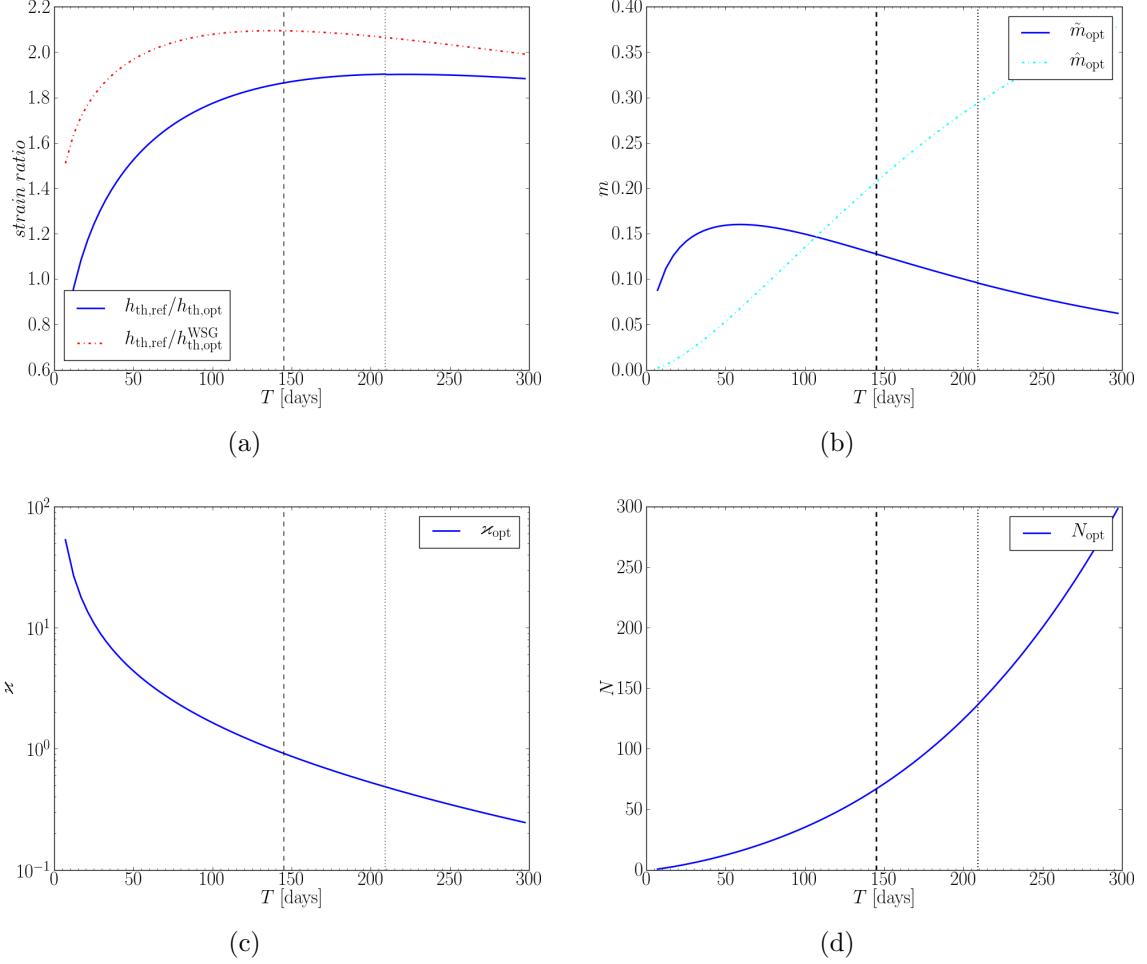


Figure 4.3.: Numerical optimal fixed- T solution for a directed CasA search as a function of T . The dashed vertical line indicates the analytical WSG-optimal solution of Eq. (4.118), while the dotted vertical line corresponds to the exact optimal solution. Panel (a) shows the weakest detectable signal strength compared to the reference value $h_{\text{th,ref}}$ of Eq. (4.114), for the exact h_{th} and for the WSG-approximated $h_{\text{th}}^{\text{WSG}}$ (using $w = 1$). (b) shows the optimal mismatch parameters $\tilde{m}_{\text{opt}}(T)$ and $\hat{m}_{\text{opt}}(T)$, (c) shows the optimal computing-cost ratio $\varkappa_{\text{opt}}(T)$ and (d) the optimal number of segments $N_{\text{opt}}(T)$ (treated as continuous).

that the search would be bounded ($\hat{a} < 0$). We can therefore use Eq. (4.103) to obtain the optimal computing-cost ratio as

$$\varkappa_{\text{opt}} = 1. \quad (4.117)$$

Note that when $w > 1.25$ we would have $\hat{a} > 0$ and therefore this search would become unbounded. From Eq. (4.78) we obtain $\tilde{m}_{\text{opt}}^{(0)} \approx 0.29$, $\hat{m}_{\text{opt}}^{(0)} \approx 0.55$, and using Eq. (4.94) we find the respective optimal mismatches as

$$\tilde{m}_{\text{opt}} \approx 0.16, \quad \hat{m}_{\text{opt}} \approx 0.24. \quad (4.118)$$

Using Eq. (4.63) we can extract the computing-cost coefficients $\tilde{\kappa} \approx 3.14 \times 10^{-17}$ and $\hat{\kappa} \approx 2.38 \times 10^{-33}$ (with time measured in seconds), and plugging this into Eqs. (4.86) and (4.87) we find the optimal StackSlide parameters as ⁸

$$\begin{aligned} N_{\text{opt}} &= 61.8, \quad \Delta T_{\text{opt}} \approx 2.3 \text{ days}, \\ T_{\text{opt}} &\approx 144.8 \text{ days}, \end{aligned} \quad (4.119)$$

which is self-consistent with the initially-assumed template-bank dimensions, as it falls into the assumed ranges for ΔT and T .

We can estimate the resulting sensitivity by solving Eqs. (4.25) and (4.26), which yields $\rho_{\Sigma}^* \approx 14.7$, and substituting into Eq. (4.29) we find a weakest detectable signal strength h_{th} of

$$\left. \frac{h_{\text{th}}}{\sqrt{S_{\text{n}}}} \right|_{\text{opt}} \approx 2.8 \times 10^{-3} \sqrt{\text{Hz}}, \quad (4.120)$$

which is an improvement on the optimal coherent sensitivity by a factor close to two.

Fig. 4.3 illustrates the behaviour of the optimal solution as a function of T without using the WSG approximation. This is obtained by numerically solving Eq. (4.87) for $\varkappa_{\text{opt}}(T)$, which yields $m_{\text{opt}}(T)$ via Eq. (4.94) and $N_{\text{opt}}(T)$ via Eq. (4.86). We see that the non-WSG approximated optimal solution results in somewhat different StackSlide parameters than the WSG solution of Eq. (4.119), but it hardly gains any further sensitivity.

Increasing the total computing cost C_0 would increase the relative advantage of the StackSlide method compared to a fully coherent search: the coherent search would gain sensitivity as $h_{\text{th}}^{-1} \propto C_0^{1/14}$ according to Eq. (4.71), while the StackSlide search would gain sensitivity as $h_{\text{th}}^{-1} \propto C_0^{1/10}$ according to Eq. (4.104) (in the WSG approximation), so here the StackSlide search is more “efficient” at converting increases of computing-power into gains of sensitivity.

⁸In this chapter we use the full refinement factor expression instead of leading order approximation, therefore the optimal parameters differ from the one found in [98].

4.4.2. All-sky CW search using Einstein@Home

As an example for a wide parameter space all-sky search with massive computing power, we consider two recent CW searches performed on the Einstein@Home computing platform [57, 58, 79], namely the StackSlide searches labelled 'S5GC1' and 'S6Bucket', which employed an efficient grid mapping implementation described in [92].

An Einstein@Home search divides the total workload into many small *workunits*, each of which covers a small fraction of the parameter space and requires only a couple of hours to complete on a host machine. These searches consisted of roughly 10^7 workunits. The E@H platform delivers a computing power of order 100 Tflop/s, and these searches ran for about 6 months each, so we can estimate their total respective computing cost is of order $C_{\text{tot}} \sim 10^{21}$ flop (i.e. ~ 1 Zeta flop). Each E@H workunit is designed to require about the same computing cost, which allows us to base the present analysis on just a single workunit.

	$\tilde{\delta}$	$\hat{\delta}$	$\hat{\eta}$	w	N	$\Delta T[\text{d}]$	$\tilde{m}$	$\hat{m}$	$C_0[\text{h}]$	κ	$\frac{h_{\text{th}}}{\sqrt{S_n}}[\sqrt{\text{Hz}}]$
S5GC1	8.7	7.7	2.0	1.1	205	1.0	0.50	0.50	0.91	2.545	2.69×10^{-3}
$T_{\text{max}} = 1 \text{ y}$	10.0	9.0	2.0	1.1	528	0.7	0.14	0.17	0.91	0.869	2.19×10^{-3}
S6Bucket	4.6	3.6	2.0	1.2	90	2.5	0.50	0.50	2.54	13.914	2.20×10^{-3}
$T_{\text{max}} = 1 \text{ y}$	3.7	2.7	2.0	1.2	175	2.1	0.58	0.32	2.54	1.815	1.93×10^{-3}

Table 4.2.: Einstein@Home example setups 'S5GC1' and 'S6Bucket', with corresponding results from an iterative optimization at fixed computing power C_0 , with assumed maximal observation time of $T_{\text{max}} = 1 \text{ y}$. The gains in weakest detectable signal strength h_{th} are $\sim 23\%$ and $\sim 14\%$, respectively.

The detector data used in these searches contained non-stationarities and gaps, and the template banks were constructed in somewhat semi-empirical ways that are hard to model analytically. In order to simplify this analysis we assume gapless stationary Gaussian data, and we use the analytic metric expressions from [56] to estimate the number of templates. This example is therefore “inspired by” recent E@H searches, but does not represent a detailed description of their computing cost or sensitivity.

The two searches 'S5GC1' and 'S6Bucket' covered a fixed spindown-range corresponding to a spindown age of $\tau = f_0/\dot{f} = 600 \text{ y}$ at a reference frequency of $f_0 = 50 \text{ Hz}$. Each workunit covers a frequency-band of $\Delta f = 0.05 \text{ Hz}$, the spindown range of $\Delta \dot{f} = 2.7 \times 10^{-9} \text{ Hz/s}$ and a (frequency-dependent) fraction q of the sky. We can incorporate the sky-fraction q by using template counts $q\mathcal{N}$ in the computing-cost expressions, where $\mathcal{N}$ are the all-sky expressions from [56]. For simplicity we fix the parameter-space dimension to $n = 4$, namely {sky, frequency, spindown}, and we use [Eq. (56),(50) in [56]]⁹ for the number of coarse-grid templates $\tilde{\mathcal{N}}$ and the refinement factor $\gamma(N)$ of [Eq. (77) in [56]] (assuming gapless data).

For a workunit at 50 Hz, the reference StackSlide parameters are:

⁹There are missing terms in both [Eqs. (57) and (83) in [56]], but one can use their Eqs. (50) and (84) instead to compute $\det g$.

- S5GC1: $q = 1/3$, $N = 205$, $\Delta T = 25$ h
- S6Bucket: $q = 1/51$, $N = 90$, $\Delta T = 60$ h

For both searches the mismatch distributions of the coarse- and fine-grid template banks are not well quantified, so we simply assume hyper-cubic template banks ($\xi = 1/3$) with $\tilde{m} = \hat{m} = 0.5$, i.e. an average total mismatch of $\langle \mu \rangle = 1/3$. Plugging these parameters into the template-counting formulae of [56], together with the timing constants of Eq. (4.115) from the Cas-A example, we find a reference per-workunit computing cost of $C_0 \approx 0.91$ h for S5GC1, and $C_0 \approx 2.5$ h for S6Bucket¹⁰. Table 4.2 shows the estimated sensitivity for these reference searches assuming the same false-alarm and false-dismissal probabilities as in the previous section.

We can apply the analytical optimal solution from Sec. 4.3 with the extracted power-law coefficients at the reference StackSlide parameters found in Table 4.2. This initially places us into the *unbounded* regime (i.e. $\hat{a} > 0$) for both ‘S5GC1’ and ‘S6Bucket’. We therefore expect to improve sensitivity by increasing T until we hit the assumed upper bound of $T_{\max} = 1$ y, so we solve Eq. (4.87) for $\varkappa_{\text{opt}}(T_{\max})$, substitute into Eq. (4.94) for $\{\tilde{m}_{\text{opt}}, \hat{m}_{\text{opt}}\}$ and obtain N_{opt} from Eq. (4.86).

In order to find a *self-consistent* solution, we need to iterate this procedure: we extract new power-law coefficients at the new solution, then re-solve until the parameters converge to better than 1% accuracy. In the case of the ‘S5GC1’ search, the converged solution falls into the unbounded regime. In the case of ‘S6Bucket’ the converged solution falls into the bounded regime, but with $T_{\text{opt}}^{(0)} > T_{\max}$. The optimal observation time is therefore $T_{\text{opt}} = 1$ y in both cases, and the resulting converged solutions and power-law coefficients are given in Table 4.2. We see that (under the present idealized conditions) we could gain $\sim 23\%$ in detectable signal strength h_{th} compared to the ‘S5GC1’ setup, and $\sim 14\%$ compared to the ‘S6Bucket’ setup.

4.4.3. All-sky search examples from CGK

The all-sky search examples studied in CGK [88] provide another interesting test case for our optimization framework. CGK considered a multi-stage optimization, but we can treat their first-stage result as a single-stage optimization problem at fixed given computing cost. CGK discussed four different cases, namely a search for either “young” (Y) neutron stars ($\tau = f/\dot{f} = 40$ y) or “old” (O) neutron stars ($\tau = 10^6$ y), using either a “fresh-data” (f) or “data-recycling” (r) mode (a distinction that is irrelevant for our present purpose). The optimized CGK StackSlide parameters and computing-cost constraints are found in [Tables I-VIII in [88]], and are summarized in Table 4.3. For the sensitivity estimates we use the same false-alarm and false-dismissal probabilities as in Sec. 4.4.1.

¹⁰The actual E@H workunits take about 6 h to complete on a machine with these timings, but these setups included bigger refinement factors γ due to gaps in the data, and used rather different template-bank designs.

Note that we expect our results to improve on the sensitivity of the CGK solution, as they incorporated an *ad-hoc* constraint of $m = \tilde{m} = \hat{m}$, and the total average mismatch in [Eq.(46) in CGK] incorrectly included only the contribution from one template grid instead of both, as discussed in Sec. 4.2.2.

The functional form of the template-bank equations (originally from BC [64]) in the CGK computing-cost model [Eq.(53) in CGK] is not consistent with the generic form of Eq. (4.43) with respect to the mismatch scaling. We therefore resort to extracting (potentially fractional) “mismatch dimensions” $\{\tilde{n}, \hat{n}\}$ using Eq. (4.64), in order to fully reproduce their computing-cost function with the power-law model of Eq. (4.61). The scaling parameters $\{\delta, \eta\}$ are extracted via Eq. (4.62) and w from Eq. (4.41). The resulting values are given in Table 4.3, assuming the FFT/resampling method for the $\mathcal{F}$ -statistic calculations.

	$\tilde{n}$	$\hat{n}$	$\tilde{\delta}$	$\hat{\delta}$	$\hat{\eta}$	w	N	$\Delta T[\text{d}]$	$\tilde{m}$	$\hat{m}$	$C_0[\text{Zf}]$	$\varkappa$	$\frac{h_{\text{th}}}{\sqrt{S_n}}[\sqrt{\text{Hz}}]$
Y/r	3.0	4.0	3.1	6.0	4.2	1.7	10	2.6	0.78	0.78	0.94	0.071	7.31×10^{-3}
$N = N_{\text{CGK}}$	3.0	4.0	3.1	6.0	4.2	1.7	10	2.1	0.17	0.48	0.94	0.482	6.38×10^{-3}
Y/f	3.0	4.0	3.1	6.0	4.2	1.7	9	2.7	0.78	0.78	0.82	0.086	7.42×10^{-3}
$N = N_{\text{CGK}}$	3.0	4.0	3.1	6.0	4.2	1.7	9	2.2	0.18	0.46	0.82	0.526	6.50×10^{-3}
O/r	2.5	2.5	3.0	10.0	7.4	1.8	8	14.8	0.35	0.35	0.74	0.004	2.61×10^{-3}
$N = N_{\text{CGK}}$	2.8	2.5	3.0	10.0	7.4	1.8	8	14.6	0.03	0.34	0.74	0.090	2.46×10^{-3}
O/f	2.5	2.6	3.0	10.0	7.3	1.7	9	11.8	0.21	0.21	0.35	0.009	2.65×10^{-3}
$N = N_{\text{CGK}}$	2.7	2.5	3.0	10.0	7.3	1.7	9	12.4	0.04	0.33	0.35	0.107	2.56×10^{-3}

Table 4.3.: CGK example search setups for young (‘Y’) and old pulsars (‘O’), using either fresh (‘f’) or recycling (‘r’) data-modes. The first line of each example gives the original CGK solution with respective extracted power-law coefficients, and the second line shows our optimal self-consistent solution with constraint $N = N_{\text{CGK}}$. The computing cost C_0 is measured in Zeta-flop ($1\text{Zf} = 10^{21}\text{flop}$).

Using the extracted scaling coefficients to compute the optimal solution from Sec. 4.3 results in a solution that is inconsistent with the initially extracted scaling coefficients. An iteration over solutions, allowing both N and T to vary, did not converge. We therefore solve a simpler problem by fixing the number of segments to the original CGK values, i.e. we constrain the solutions to $N = N_{\text{CGK}}$. We proceed by solving Eq. (4.86) for $\varkappa_{\text{opt}}(N)$, closing the solution via Eqs. (4.94) and (4.87). We then extract new power-law coefficients at this solution and iterate this procedure until convergence to better than 1% accuracy is achieved. The resulting fixed- N optimal solutions are given in Table 4.3. The respective improvements of the weakest detectable signal strength h_{th} compared to the original CGK solutions are $\sim 15\%$ in the young (Y) pulsar case, and $\sim 5\%$ in the old (O) pulsar case.

4.4.4. CWs from binary neutron stars

For CWs from NSs in binary systems with known sky-position (such as Sco-X1 and other LMXBs), the search parameter space typically consists of the intrinsic signal frequency and orbital parameters of the binary system, i.e. (projected) semi-major axis,

orbital period P , periape angle, eccentricity and eccentric anomaly. The corresponding template-counting formulae were initially studied in [99] for coherent searches. These have recently been extended to semicoherent searches by Messenger [93], giving explicit template scalings in two limiting cases, namely (i) *short* coherent segments compared to the orbital period, i.e. $\Delta T \ll P$, and (ii) *long* coherent segments, i.e. $\Delta T \gg P$.

(i) Short coherent segments ($\Delta T \ll P$)

One can change parameter-space coordinates and Taylor-expand in small $\Delta T/P \ll 1$ to obtain the coherent template scaling [Eq. (24) in [93]]:

$$\tilde{\mathcal{N}}_{\tilde{n}} \propto \Delta T^{\tilde{n}(\tilde{n}+1)/2}, \quad (4.121)$$

where $\tilde{n}$ is the effective coherent parameter-space dimension using the new coordinates. The coherent cost power-law coefficients are therefore $\tilde{\delta} = \tilde{n}(\tilde{n} + 1)/2 + \Delta\tilde{\delta}$ and $\tilde{\eta} = 1$.

The semicoherent template scaling including eccentricity results in a 6-dimensional semicoherent template bank, i.e. $\hat{n} = 6$, and a template scaling [Eq. (28) in [93]] of $\hat{\mathcal{N}}_{\text{ecc}} \propto N \Delta T^7$. In the case of small eccentricity one has $\hat{n} = 4$, and the template scaling given in [Eq. (29) in [93]] is $\hat{\mathcal{N}}_{\text{circ}} \propto N \Delta T^5$. In both cases the semicoherent power-law exponents satisfy $\hat{\delta} \geq 5$, and $\hat{\eta} = 2$, resulting in the critical parameter $\hat{a} > 0$. This implies that the boundedness-condition Eq. (4.102) is always violated, i.e. one should use all the available data.

(ii) Long coherent segments ($\Delta T \gg P$)

In this limit the template scalings in both the coherent and semicoherent step are [Eqs. (32,33) in [93]]: $\tilde{\mathcal{N}} \propto \hat{\mathcal{N}} \propto \Delta T^2$, which is unusual as there is *no refinement*. Therefore $\tilde{\eta} = \hat{\eta} = 1$, and $\tilde{\delta} = 2 + \Delta\tilde{\delta}$, while $\hat{\delta} = 2$, and therefore $\hat{\varepsilon} = 1$. We see that always $\tilde{a} > 0$ and $\hat{a} = 2(w - 1) > 0$, and therefore binary-CW searches in the long-segment limit also fall into the *unbounded* regime, i.e. one should use all the data.

4.5. Discussion

In this chapter we have derived an improved estimate of the StackSlide sensitivity scaling, correctly accounting for the mismatches from both coarse- and fine-grid template banks, which had been overlooked by previous studies. By locally fitting sensitivity and computing-cost functions to power laws we are able to derive fully analytical self-consistency relations for the optimal sensitivity at fixed computing cost. This solution separates two different regimes depending on the critical parameter $\hat{a}$ of Eq. (4.102): a bounded regime with a finite optimal $T_{\text{opt}}^{(0)}$, and an unbounded regime where $T_{\text{opt}}^{(0)} \rightarrow \infty$.

Several practical examples are discussed in order to illustrate the application of this framework. The corresponding sensitivity gains in terms of the weakest detectable signal strength h_{th} are found to be $\sim 100\%$ compared to a fully coherent directed search

for CasA, and about 5% – 20% compared to previous StackSlide searches such as Einstein@Home and the examples given in CGK [88]. We show that CW searches for binary neutron stars seem to generally fall into the unbounded regime where all the available data should be used irrespective of available computing power.

This study only considered single-stage StackSlide searches on Gaussian stationary gapless data from detectors with identical noise-floors. In the next chapter we extend this analysis to more realistic data conditions.

5. Optimal StackSlide method at fixed computing cost using real data

The enormous computational requirement of the wide parameter-space searches for continuous gravitational waves impose a cautious use of the available computing resources, i.e., we always aim at maximal sensitivity. In this respect to maximize the sensitivity of a semicoherent search, for example StackSlide, means that we need to choose the optimal search parameters, namely number of segments, segment duration and optimal maximal mismatch on the coarse and fine grid. How to do this at fixed computing cost and under the ideal conditions of constant noise floor and data without gaps has been studied in Chap. 4, where we derived analytical expressions to determine the optimal search parameters. However under real conditions the available data can be fragmented, e.g., due to down-time of the detectors, as well as the noise floor may be changing. The fragmentation of the data can significantly affect the computing cost function and thus the optimal search parameters. On the other hand, the noise fluctuations suggest the use of a data selection procedure in order to spend the available computing cycles searching data of higher quality.

In the following we extend Chap. 4 to these more realistic conditions by taking into account possible gaps in the data and noise level changes. We show, how the real conditions manifest in the sensitivity function, but nevertheless it is possible to decouple the related data selection problem from the sensitivity maximization. Therefore we can apply a numerical optimization to find optimal semicoherent search parameters. We also describe a suitable data selection algorithm. The outcome of the proposed numerical optimization are the optimal search parameters and the selected data, thus the search can be performed in practice. We first test the numerical optimization procedure under ideal conditions and obtain the results of the analytical method. Then we give examples of practical application.

This chapter is organized as follows. In Sec. 5.1 we introduce the ingredients of the search-optimization method, i.e., the threshold signal-to-noise ratio (SNR), the sensitivity function, and the computing cost function. The numerical optimization of the search parameters and in particular the data selection method are described in Sec. 5.2. In Sec. 5.3 we give examples of practical application and discuss in Sec. 5.4.

Notation

We follow the notation of Chap. 4, i.e., we use tilde when referring to fully coherent quantities, $\tilde{Q}$, and overhat when referring to semicoherent quantities, $\hat{Q}$.

5.1. Threshold SNR, sensitivity function and computing cost

In this section we introduce the main ingredients needed to define the optimization problem, i.e., to find the number of segments N with segment duration ΔT and coarse and fine grid mismatch, $\tilde{m}$ resp. $\hat{m}$, which maximize the sensitivity of the search at fixed computing cost C_0 . These ingredients are the threshold SNR required for detection, the sensitivity function, which we want to maximize and the computing cost function.

5.1.1. Threshold SNR

A claim for detection, when searching for a signal, in particular a weak signal, in the presence of noise is sensible only in the context of two well defined quantities. The first one, called false-alarm probability, is the probability to falsely claim a detection when the signal is not present in the data. The second quantity, called false-dismissal probability, is the probability to miss the signal even if the signal is indeed present in the data.

When a signal is present in the data the StackSlide statistic follows a noncentral χ^2 distribution with $4N$ degrees of freedom. Using the notation and definitions of the previous chapter, we denote this by $\chi_{4N}^2(\Sigma, \rho_\Sigma^2)$, where Σ is the StackSlide statistic, and ρ_Σ^2 is the noncentrality parameter, i.e., the sum of the squared signal-to-noise (SNR) ratio of the individual segments $\rho_\Sigma^2 \equiv \sum_{i=1}^N \rho_i^2$. We can integrate the false-alarm probability

$$p_{\text{fA}}(\Sigma_{\text{th}}) = \int_{\Sigma_{\text{th}}}^{\infty} d\Sigma \chi_{4N}^2(\Sigma; 0) , \quad (5.1)$$

and by inversion for a given false-alarm p_{fA}^* obtain the threshold Σ_{th} . For a pre-defined false-dismissal p_{fD}^* probability

$$p_{\text{fD}}(\Sigma_{\text{th}}, \rho_\Sigma^2) = \int_{-\infty}^{\Sigma_{\text{th}}} d\Sigma \chi_{4N}^2(\Sigma; \rho_\Sigma^2) , \quad (5.2)$$

using Σ_{th} we want to obtain the critical non-centrality

$$\rho_\Sigma^{*2} = \rho_\Sigma^2(p_{\text{fA}}^*, p_{\text{fD}}^*, N) , \quad (5.3)$$

and thus the threshold SNR ρ_Σ^* .

The computation of the critical non-centrality ρ_Σ^{*2} is complicated by the fact that for wide parameter-space searches the right-hand side of Eq. (5.2) requires averaging over sky position and polarization parameters of the signal, but still at fixed h_0 . In the previous chapter we assumed a signal population of constant SNR. Therefore by application of the central limit theorem and approximation of the χ^2 distribution by a Gaussian distribution, we analytically integrated and inverted Eq. (5.2) to obtain (5.3).

For weak signals and large number of segments this resulted in the “weak-signal Gauss (WSG) approximation”, Eq. (4.38), for the critical non-centrality parameter

$$\rho_{\Sigma}^{*2}(p_{\text{fA}}^*, p_{\text{fD}}^*, N) \approx 2\sqrt{4N}(\text{erfc}^{-1}(2p_{\text{fA}}^*) + \text{erfc}^{-1}(2p_{\text{fD}}^*)) . \quad (5.4)$$

With this we define the per-segment threshold SNR

$$\rho^* \equiv \sqrt{\frac{\rho_{\Sigma}^{*2}(p_{\text{fA}}^*, p_{\text{fD}}^*, N)}{N}} . \quad (5.5)$$

Recently a new semi-analytical method has been developed to estimate the sensitivity of a search [100]. The accuracy of this technique is within the calibration error of the gravitational wave detector with results similar to the sensitivity estimates performed with Monte-Carlo methods. The assumption of signal distribution of constant SNR has been relaxed in [100], where a semi-analytical approximation for the computation of an isotropic threshold SNR has been introduced. We refer to this method as the KWS approximation.

In the KWS approximation the averaged over segments threshold ρ^* is obtained recursively. At iteration i the value of ρ^* is

$$\rho = F(\rho) , \quad (5.6)$$

where

$$\rho_{i+1} = F\left(\frac{\rho_i + \rho_{i-1}}{2}\right) . \quad (5.7)$$

For the exact form of F and the details required to implement the method in practice see [100].

For comparison of the WSG and KWS approximations in Fig. 5.1 we plot the ratio $\rho_{\text{WSG}}^*/\rho_{\text{KWS}}^*$ at fixed $p_{\text{fA}} = 1 \times 10^{-10}$ and $p_{\text{fD}} = 0.1$. The threshold SNR is underestimated in the WSG approximation, thus we underestimate the detectable intrinsic amplitude h_0 , i.e., we overestimate the sensitivity of the search; see Eq. (5.13). This effect is stronger for small number of segments, where the WSG approximation breaks. In this chapter we use the WSG approximation only in one example, where we determine optimal semicoherent search parameters analytically for comparison with the result of the numerical method proposed in Sec. 5.3.

5.1.2. Sensitivity function

The signal strength h_{rms} in the detector, depends on the intrinsic amplitude h_0 , the sky-position of the source, the polarization angles and the detector orientation. Averaging isotropically over the sky-position and polarization yields

$$\langle h_{\text{rms}}^2 \rangle_{\text{sky, pol}} = \frac{2}{25} h_0^2 . \quad (5.8)$$

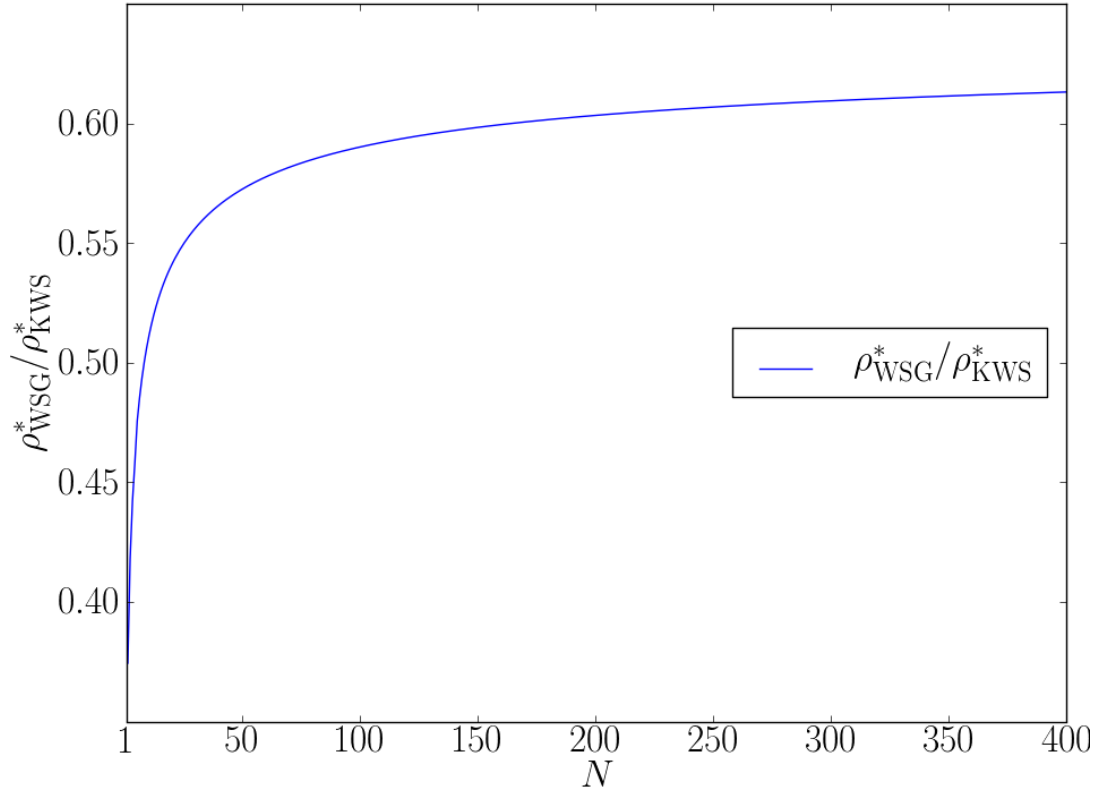


Figure 5.1.: Ratio of threshold SNR computed with the WSG and KWS approximation as function of the number of segments N at fixed false-alarm $p_{\text{fA}} = 1 \times 10^{-10}$ and false-dismissal $p_{\text{fD}} = 0.1$ probability. The threshold SNR is underestimated in the WSG approximation, leading to underestimate of the intrinsic amplitude h_0 , and therefore overestimate of the sensitivity of the search.

Under ideal conditions of data without gaps with duration T and constant noise floor S_n the accumulated SNR² in a semicoherent search is (see Eq. (4.24)):

$$\rho_\Sigma^2 = [1 - \xi(\tilde{m} + \hat{m})] \frac{2N_{\text{det}}}{S_n} h_{\text{rms}}^2 T , \quad (5.9)$$

where $\xi \in (0, 1)$ is the geometrical factor used to estimate the average mismatch of the template bank, $\tilde{m}$ is the mismatch on the coarse grid, $\hat{m}$ is the mismatch on the fine grid and N_{det} is the number of detectors. To take into account possible noise floor fluctuations, with the data in form of N_{SFT} number of Short Fourier Transforms (SFTs) of duration T_{SFT} , the noise floor can be written as, Eq. (2.81),

$$\mathcal{S}^{-1}(f_0) \equiv N_{\text{SFT}}^{-1} \sum_{n=1}^{N_{\text{SFT}}} S_n^{-1}(f_0) , \quad (5.10)$$

where S_n is the per SFT noise Power Spectral Density (PSD) estimated at frequency f_0 . To account for the possible gaps in the data we define the actual amount of available data

$$T_{\text{data}} \equiv N_{\text{SFT}} T_{\text{SFT}} . \quad (5.11)$$

Using Eqs. (5.8), (5.10) and (5.11) we can rewrite Eq. (5.9) to obtain the accumulated SNR² in a semicoherent search under real conditions, namely

$$\langle \rho_\Sigma^2 \rangle_{\text{sky, pol}} = \frac{4}{25} [1 - \xi(\tilde{m} + \hat{m})] h_0^2 T_{\text{data}} \mathcal{S}^{-1} . \quad (5.12)$$

To estimate the minimal detectable intrinsic amplitude h_0 at fixed false-alarm p_{fA}^* and false-dismissal p_{fD}^* probability we use the per-segment threshold SNR ρ^* . With $\langle \rho_\Sigma^2 \rangle_{\text{sky, pol}} \equiv N \rho^{*2}$ substitution in Eq. (5.12) and rearrangement yields

$$h_0 = \frac{5}{2} [1 - \xi(\tilde{m} + \hat{m})]^{-1/2} \rho^* \sqrt{N} \sqrt{\frac{1}{\mathcal{G}}} , \quad (5.13)$$

where

$$\mathcal{G} \equiv T_{\text{data}} \mathcal{S}^{-1} \quad (5.14)$$

is the *goodness* of the data. Eq. (5.13) is the function that we need to minimize under the constraint of fixed computing cost C_0 in order to maximize the sensitivity of the search.

5.1.3. Computing cost

The total computing cost $C_{\text{tot}}(\tilde{m}, \hat{m}, N, \Delta T, N_{\text{SFT}})$ of the StackSlide method is composed by the cost $\tilde{C}(\tilde{m}, \Delta T, N_{\text{SFT}})$ to compute the $\mathcal{F}$ -statistic [91, 46] on the coarse grid and the cost $\hat{C}(\hat{m}, N, \Delta T)$ to sum these $\mathcal{F}$ -statistic values across all segments on the fine grid, thus

$$C_{\text{tot}}(\tilde{m}, \hat{m}, N, \Delta T, N_{\text{SFT}}) = \tilde{C}(\tilde{m}, \Delta T, N_{\text{SFT}}) + \hat{C}(\hat{m}, N, \Delta T) . \quad (5.15)$$

The computing cost of the coherent step using the SFT method is

$$\tilde{C}(\tilde{m}, \Delta T, N_{\text{SFT}}) = \tilde{\mathcal{N}}(\tilde{m}, \Delta T) N_{\text{SFT}} \tilde{c}_0^{\text{SFT}}, \quad (5.16)$$

where N_{SFT} is the total number of used SFTs, $\tilde{\mathcal{N}}(\tilde{m}, \Delta T)$ is the number of coarse-grid templates and $\tilde{c}_0^{\text{SFT}}$ is an implementation- and hardware-dependent fundamental computing cost.

The incoherent computing cost is

$$\hat{C}(\hat{m}, N, \Delta T) = N \hat{\mathcal{N}}(\hat{m}, \Delta T, N) \hat{c}_0, \quad (5.17)$$

where $\hat{\mathcal{N}}(\hat{m}, \Delta T, N)$ is the number of fine-grid templates and $\hat{c}_0$ is the implementation- and hardware-dependent fundamental cost of adding $\mathcal{F}$ -statistic values.

5.1.4. Templates counting

The general expression for the number of templates required to cover some parameter space $\mathbb{P}$ is

$$\mathcal{N} = \theta_n m^{-n/2} \mathcal{V}_n, \quad \text{with} \quad \mathcal{V}_n \equiv \int_{\mathbb{T}_n} d^n \lambda \sqrt{\det g}, \quad (5.18)$$

where θ_n is the normalized thickness of the search grid, m is the maximal-mismatch, $\det g$ is the determinant of the parameter-space metric g_{ij} and $\mathcal{V}_n$ is the metric volume of the n -dimensional space of the template bank. For hyper-cubic lattice the normalized thickness is $\theta_{\mathbb{Z}_n} = n^{n/2} 2^{-n}$, while for an A^* lattice it is $\theta_{A_n^*} = \sqrt{n+1} \left\{ \frac{n(n+2)}{12(n+1)} \right\}^{n/2}$ [54]. The choice of the dimensionality of the template bank is subject to the maximization of the number of templates, namely

$$\tilde{\mathcal{N}}_{\tilde{n}} = \max_n \tilde{\mathcal{N}}_n, \quad \text{and} \quad \hat{\mathcal{N}}_{\hat{n}} = \max_n \hat{\mathcal{N}}_n. \quad (5.19)$$

In the previous chapter we used the factorization of the semicoherent metric volume

$$\hat{\mathcal{V}}_{\hat{n}}(N, \Delta T) = \gamma_{\hat{n}}(N) \tilde{\mathcal{V}}_{\tilde{n}}(\Delta T), \quad (5.20)$$

to derive the general power-law computing-cost model, as in the gap-less data case, the refinement factor $\gamma_{\hat{n}}(N)$ is only function of the number of segment N . However using real data introduces an additional dependency on the time span of the search through the actual position of the segments in time. For details see, e.g., [56]. We aim to model the real conditions as closely as possible, in the numerical optimization we therefore directly compute the semicoherent metric

$$\hat{g}(t_{\text{ref}}) = \frac{1}{N} \sum_{i=1}^N \tilde{g}_i(t_i, \Delta T, t_{\text{ref}}), \quad (5.21)$$

where $\tilde{g}_i(t_i, \Delta T, t_{\text{ref}})$ is the coherent metric of segment i at fixed reference time t_{ref} . To compute the coherent metric $\tilde{g}_i(t_i, \Delta T, t_{\text{ref}})$ in this chapter we use the analytical expressions found in [56].

5.1.5. The choice of spindown parameter space

The choice of the dimensionality of the template bank through Eqs. (5.19) is possible only for a simple rectangular geometry of the parameter space.

As discussed in Chap. 2, using the spindown age of a potential source [64]:

$$\tau = f/|\dot{f}| . \quad (5.22)$$

the spindown search band in dimension k is

$$|f^{(k)}| \leq k! \frac{f}{\tau^k} . \quad (5.23)$$

This however means that the spindown band is frequency dependent, which may be unpractical. Therefore if we keep the number of templates in the spindown space constant by fixing a minimal detectable spindown age τ_0 at some frequency f_0 , the detectable spindown age at frequency f yields

$$\tau(f) = \tau_0 f / f_0 . \quad (5.24)$$

This would define the simplest possible parameter-space volume for optimization, namely a “box”.

A more complicated triangular parameter-space shape, has been discussed in [62] and used in the search for gravitational waves from the supernova remnant Cassiopeia-A [63]. The parameters of a search over such space are difficult to optimize as the spindown order may vary even in infinitesimally small slices of a frequency band; see Fig. 2.3. In the following we neglect this fact in order to compare the outcome of the optimization with previously obtained results.

5.2. Numerical optimization procedure for a semicoherent StackSlide search with a fixed frequency band

In this section we consider the practical implementation of a numerical optimization procedure to find optimal search parameters for a semicoherent StackSlide search.

5.2.1. Definition of the optimization problem

To maximize the sensitivity of the search, i.e., to minimize the measurable intrinsic amplitude, we need to minimize the function given in Eq. (5.13), namely:

$$h_0(\tilde{m}, \hat{m}, N, \mathcal{G}) = \frac{5}{2} [1 - \xi(\tilde{m} + \hat{m})]^{-1/2} \rho^* \sqrt{N} \sqrt{\frac{1}{\mathcal{G}}} \quad (5.25)$$

under the constraints

$$\begin{aligned}
C_{\text{tot}} &\leq C_0, \\
0 &< \tilde{m} < 1, \\
0 &< \hat{m} < 1, \\
\tilde{N} &\leq \hat{N},
\end{aligned} \tag{5.26}$$

where C_0 is a given maximal available computing power. Note that in practice the cost constraint is difficult to be given as an equality. However a reasonable algorithm should find a solution for which C_{tot} becomes approximately C_0 . At first glance the form of Eq. (5.25) appear to be not well suited for optimization, because one of the parameters of the semicoherent search, namely the segment duration ΔT , is not explicitly present in the equation. In fact, there is an implicit dependency on ΔT through T_{data} ; the quantity needed to compute the goodness of the data $\mathcal{G}$, which we have to maximize in order to minimize h_0 . In other words, we need a data selection procedure. This decoupling of the data selection allows us in practice to minimize $h_0(\tilde{m}, \hat{m}, N, \Delta T)$, which is a 4D optimization problem over $\tilde{m}$, $\hat{m}$, N and ΔT . The dependency of h_0 on ΔT through $\mathcal{G}$ will become clearer in the next subsection, where we explain the data selection procedure.

5.2.2. Data selection

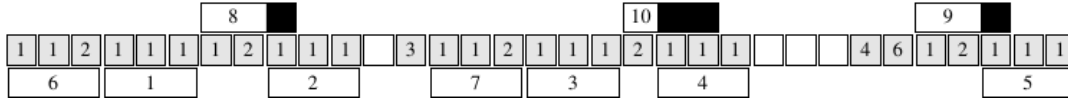


Figure 5.2.: Schematic representation of the data selection procedure for requested $N = 10$ segments with duration $\Delta T = 3$ time units. The middle line of squares in gray are the available SFTs of unit time, where the number inside the square denotes the PSD. The white squares without a number are gaps in the data. The selected segments are the white rectangles, where the number inside is the number of the segment. Note how segments 8, 9 and 10 overlap in time with their neighboring segments, however they do not share data. This is depicted by the black fill of the segments.

For a given amount of data, we need an algorithm to select the best data in order to maximize the goodness $\mathcal{G}$, i.e., pick as much data T_{data} as possible of lowest noise level $\mathcal{S}$ as possible. For a requested pair of number of segments N with duration ΔT and given set of SFTs with duration T_{SFT} , which are ordered in time by increasing timestamp t_j , the data selection consist in the following steps:

1. For each timestamp t_j find all SFTs in the time interval $[t_j, t_j + \Delta T - T_{\text{SFT}}]$ and compute $\mathcal{G}_j$.

2. Select the segment starting from t_j such that $\mathcal{G}_j$ is maximal and remove the timestamps of the SFTs, which belong to the selected segment ¹.
3. Repeat steps 1. and 2. until N segments are selected or there is no more data left.

An example of data selection is schematically presented in Fig. 5.2, where we select $N = 10$ segments of duration $\Delta T = 3$ time units out of data set with $N_{\text{SFT}} = 33$ SFTs of unit time distributed in $T = 37$ time units. Three of the selected segments overlap in time with their neighbors, however these overlapping segments do not share data, as the data is used from the neighboring segments selected first. Such partial segments suggest grid construction based on the actual length of the segment instead of the maximal length ΔT , which would reduce the coherent part of the total computing cost in these cases. However this would complicate the combination of the $\mathcal{F}$ -statistic values in the semicoherent step of the StackSlide search, thus the overall effect remains unclear. Therefore we stick to a constant grid for every segment.

Before proceeding with the examples of practical application in the next section, we note, that an equivalent data selection procedure has been used to select data for the recent all-sky Einstein@Home searches for continuous gravitational waves in LIGO S6 data.

5.3. Examples of practical application

In this section we apply the numerical optimization procedure to a directed semicoherent search, first using simulated and then real detector data (the results are summarized in Table 5.2). For comparison with the analytical solution found in the “Directed search for Cassiopeia-A” example in the previous chapter 4.4.1, we use the same search volume enclosed in the frequency band $f \in [100, 300]$ Hz with spindown ranges corresponding to a spindown-age $\tau_{\text{min}} = 300$ y. The computing cost constraint is $C_0 \approx 472$ days on a single computing core, where the fundamental computing constants are

$$\hat{c}_0^{\text{SFT}} = 7 \times 10^{-8} \text{ s}, \quad \hat{c}_0 = 6 \times 10^{-9} \text{ s} . \quad (5.27)$$

In the following we fix the false-alarm $p_{\text{FA}} = 1 \times 10^{-10}$ and false-dismissal probability $p_{\text{FD}} = 0.1$. We assume an A^* search grid, for which $\xi \approx 0.5$. We report the minimal detectable signal strength, from Eq. (5.8)

$$h_{\text{th}} = \frac{\sqrt{2}}{5} h_0 . \quad (5.28)$$

The weakest detectable signal, as estimated in the previous chapter using the constant SNR approximation with $T_{\text{data}} = 0.7 \times 2 \times 12$ days, $\xi = 0.5$ and $\tilde{m} = 0.2$ is

$$\left. \frac{h_{\text{th}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 5.2 \times 10^{-3} \sqrt{\text{Hz}} . \quad (5.29)$$

¹Working with real numbers it is unlikely to find segments with exactly equal goodness, but in such rare cases we choose the earliest segment in time.

which in the KWS approximation yields

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 7.67 \times 10^{-3} \sqrt{\text{Hz}} . \quad (5.30)$$

We perform the numerical optimization with the NOMAD [101] implementation of a Mesh Adaptive Direct Search (MADS) [70, 71, 72] algorithm for constrained derivative-free optimization. For each of the following examples we run the procedure 50 times from a common initial starting point:

$$\begin{aligned} N_0 &= 200, & \Delta T_0 &= 1 \text{ days} \\ \tilde{m}_0 &= 0.5, & \hat{m}_0 &= 0.5, \end{aligned}$$

while we use different mesh coarsening and mesh update basis parameters. These are internal parameters for the MADS algorithm controlling the evolution of the mesh if a better solution than the current one is found. We use these multiple runs of the optimization effectively to escape local extrema.

5.3.1. Directed search using simulated data

Case A: Gapless data with constant noise floor

We first consider optimization using simulated data from 2 detectors spanning 365 days, without gaps, and with a constant noise floor $\sqrt{S_n} = 1 / \sqrt{\text{Hz}}$. Using the analytical optimization method discussed in the previous chapter and the WSG approximation to obtain optimal parameters, the weakest detectable signal strength in the KWS approximation equals to

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 3.6 \times 10^{-3} \sqrt{\text{Hz}} . \quad (5.31)$$

The optimal maximal mismatch on the coarse and fine grid is

$$\tilde{m}_{\text{opt}} = 0.16, \quad \hat{m}_{\text{opt}} = 0.24, \quad (5.32)$$

and the optimal number of segments N_{opt} , segment duration ΔT_{opt} and total observation time T_{opt} are

$$\begin{aligned} N_{\text{opt}} &= 76.5, & \Delta T_{\text{opt}} &\approx 2.0 \text{ days}, \\ T_{\text{opt}} &\approx 155.5 \text{ days}. \end{aligned} \quad (5.33)$$

The results of the numerical optimization when using the WSG approximation to find optimal parameters is plotted on the left side of Fig. 5.3. We find an optimal solution yielding a detectable signal strength h_{th} in the KWS approximation

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 3.5 \times 10^{-3} \sqrt{\text{Hz}}, \quad (5.34)$$

with slightly different optimal mismatches

$$\tilde{m}_{\text{opt}} \approx 0.09, \quad \hat{m}_{\text{opt}} \approx 0.29, \quad (5.35)$$

and segment number and duration, of

$$\begin{aligned} N_{\text{opt}} &= 185.0, \quad \Delta T_{\text{opt}} \approx 1.3 \text{ days}, \\ T_{\text{opt}} &\approx 231.3 \text{ days}, \end{aligned} \quad (5.36)$$

compared to the analytical solution.

The outcome of the optimization when using the KWS approximation is plotted on the right side of Fig. 5.3. The detectable signal strength in this case is

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 3.5 \times 10^{-3} \sqrt{\text{Hz}}, \quad (5.37)$$

with optimal mismatches

$$\tilde{m}_{\text{opt}} \approx 0.10, \quad \hat{m}_{\text{opt}} \approx 0.29, \quad (5.38)$$

and segment number and duration

$$\begin{aligned} N_{\text{opt}} &= 188, \quad \Delta T_{\text{opt}} \approx 1.3 \text{ days}, \\ T_{\text{opt}} &\approx 235 \text{ days}. \end{aligned} \quad (5.39)$$

The gain in sensitivity of this toy semicoherent search compared to the original fully coherent search using the more accurate KWS approximation is ≈ 2.2 .

Case B: Data with gaps and constant noise floor

We consider now data with gaps allowing a duty cycle (fraction of actually available data T_{data} in a given time span T_{span} , $\epsilon \equiv T_{\text{data}}/T_{\text{span}}$) of 70 % per detector, while the noise floor is still constant $\sqrt{S_n} = 1 / \sqrt{\text{Hz}}$. The results of the numerical optimization when using the KWS approximation to compute the threshold SNR are plotted in Fig. 5.4. The detectable signal strength in this case is

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 4.4 \times 10^{-3} \sqrt{\text{Hz}}, \quad (5.40)$$

with optimal mismatches

$$\tilde{m}_{\text{opt}} \approx 0.04, \quad \hat{m}_{\text{opt}} \approx 0.46, \quad (5.41)$$

and segment number and duration

$$\begin{aligned} N_{\text{opt}} &\approx 364, \quad \Delta T_{\text{opt}} \approx 0.8 \text{ days}, \\ T_{\text{opt}} &\approx 364.9 \text{ days}. \end{aligned} \quad (5.42)$$

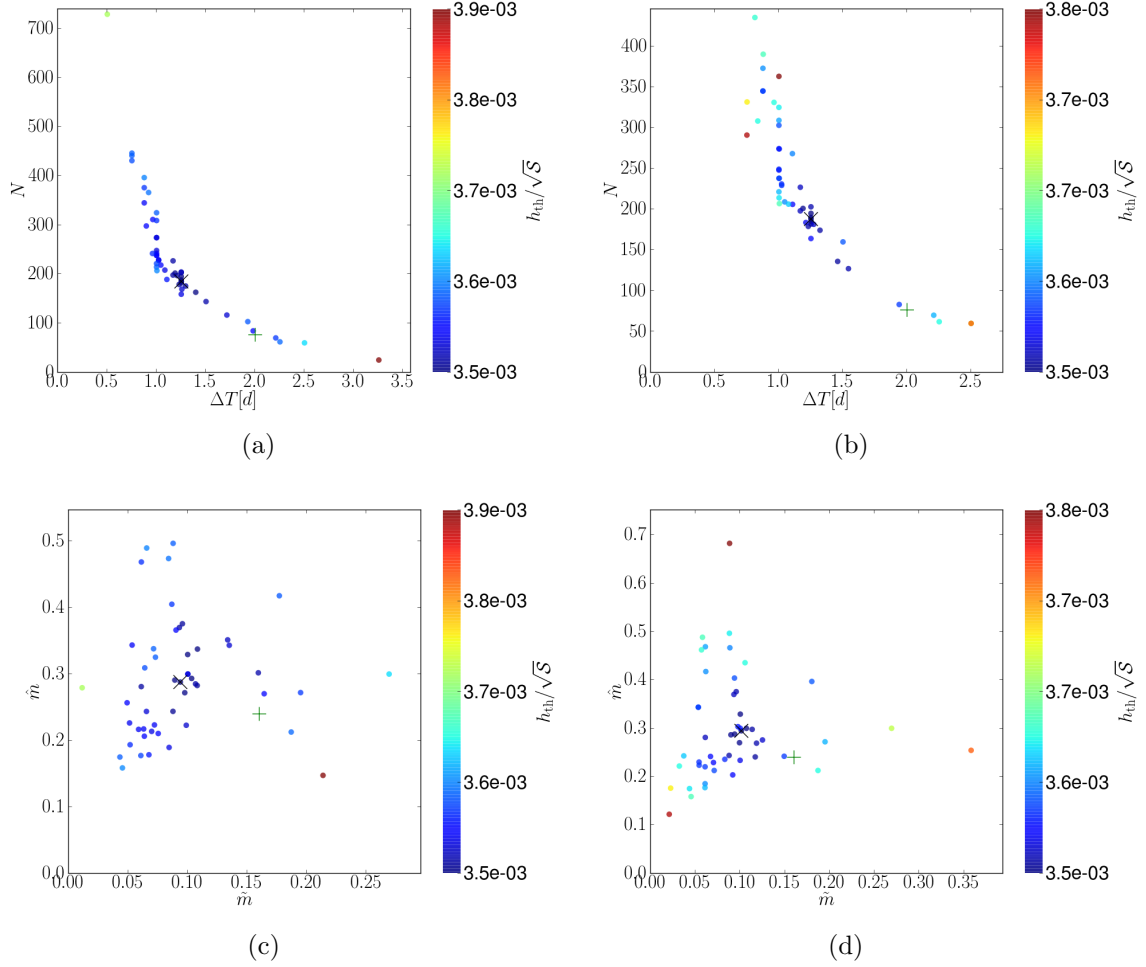


Figure 5.3.: Semicoherent search optimization using the WSG approximation (left) and KWS approximation (right) for data from 2 detectors, without gaps and of constant noise floor $\sqrt{S_n} = 1/\sqrt{\text{Hz}}$. The sensitivity is always given in the KWS approximation. The cost constraint is $C_0 = 472.0$ days. The best numerical solution is denoted with $\times$ and the optimal analytical solution with $+$. The most sensitive WSG optimal solution is: $\tilde{m}_{\text{opt}} \approx 0.09$, $\hat{m}_{\text{opt}} \approx 0.29$, $\Delta T_{\text{opt}} \approx 1.3$ days, $N_{\text{opt}} \approx 185.0$, $T_{\text{opt}} \approx 231.3$ days, $\tilde{n} = 2$, $\hat{n} = 3$, $h_{\text{th}}/\sqrt{S} \approx 3.5 \times 10^{-3} \sqrt{\text{Hz}}$. The most sensitive KWS optimal solution is: $\tilde{m}_{\text{opt}} \approx 0.10$, $\hat{m}_{\text{opt}} \approx 0.29$, $\Delta T_{\text{opt}} \approx 1.3$ days, $N_{\text{opt}} \approx 188.0$, $T_{\text{opt}} \approx 235.0$ days, $\tilde{n} = 2$, $\hat{n} = 3$, $h_{\text{th}}/\sqrt{S} \approx 3.5 \times 10^{-3} \sqrt{\text{Hz}}$. (a) WSG optimal total observation time T and number of segments N . (b) KWS optimal total observation time T and number of segments N . (c) WSG optimal maximal mismatch on the coarse and fine grid, $\tilde{m}$ resp. $\hat{m}$. (d) KWS optimal maximal mismatch on the coarse and fine grid, $\tilde{m}$ resp. $\hat{m}$.

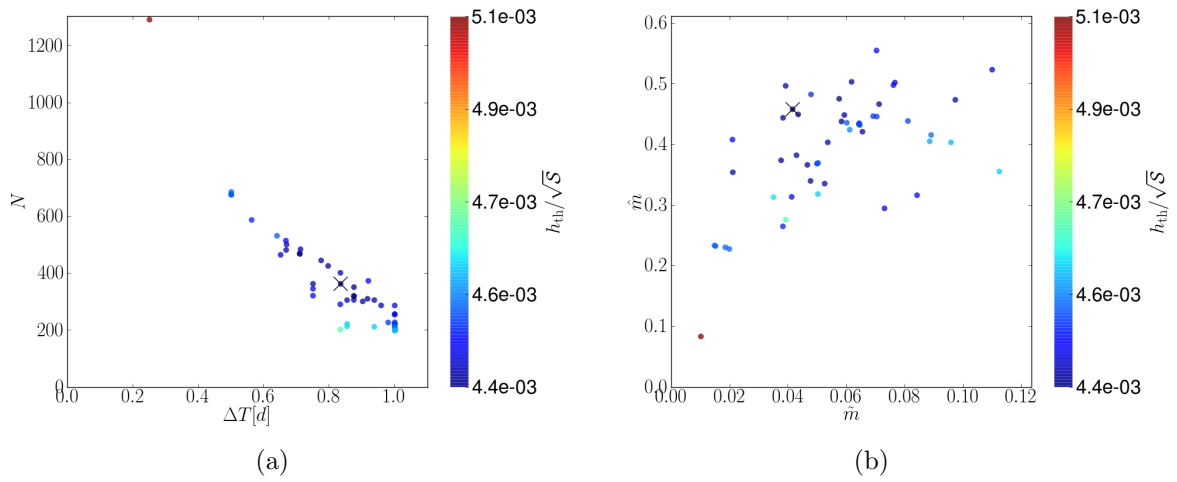


Figure 5.4.: Semicoherent search optimization using the KWS approximation for data from 2 detectors with duty cycle of 70 % and of constant noise floor $\sqrt{S_n} = 1/\sqrt{\text{Hz}}$. The cost constraint is $C_0 = 472.0$ days. The best numerical solution is denoted with $\times$. The parameter values are: $\tilde{m}_{\text{opt}} \approx 0.04$, $\hat{m}_{\text{opt}} \approx 0.46$, $\Delta T_{\text{opt}} \approx 0.8$ days, $N_{\text{opt}} \approx 364.2$, $T_{\text{opt}} \approx 364.9$ days, $\tilde{n} = 2$, $\hat{n} = 3$, $h_{\text{th}}/\sqrt{S} \approx 4.4 \times 10^{-3} \sqrt{\text{Hz}}$. (a) KWS optimal total observation time T and number of segments N . (b) KWS optimal maximal mismatch on the coarse and fine grid, $\tilde{m}$ resp. $\hat{m}$.

Case C: Data with gaps and noise floor fluctuations

In this example we further relax the requirements on the data by allowing noise floor fluctuations, while keeping the duty cycle of 70 % per detector. For each SFT the PSD has been drawn from a Gaussian distribution with mean $E[\sqrt{S_n}] = 1 / \sqrt{\text{Hz}}$ and standard deviation $\sigma[\sqrt{S_n}] = 15 \times 10^{-2} \text{ Hz}^{-1/2}$. The outcome of the optimization when using the KWS approximation is plotted on the right side of Fig. 5.4. The detectable signal strength in this case is

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{S}} \right|_{\text{opt}} \approx 4.4 \times 10^{-3} \sqrt{\text{Hz}} , \quad (5.43)$$

with optimal mismatches

$$\tilde{m}_{\text{opt}} \approx 0.05 , \quad \hat{m}_{\text{opt}} \approx 0.48 , \quad (5.44)$$

and segment number and duration

$$\begin{aligned} N_{\text{opt}} &\approx 325 , \quad \Delta T_{\text{opt}} \approx 0.9 \text{ days} , \\ T_{\text{opt}} &\approx 364.5 \text{ days} . \end{aligned} \quad (5.45)$$

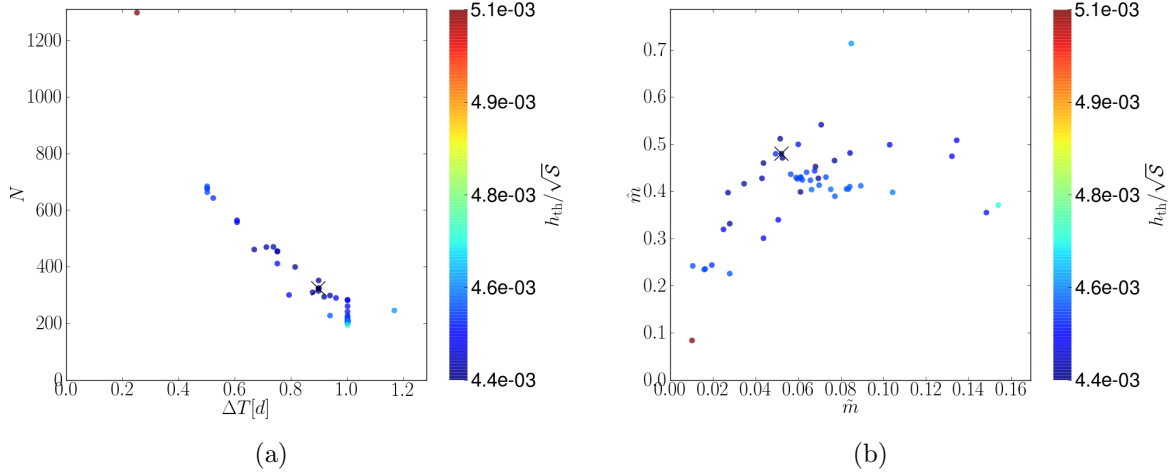


Figure 5.5.: Semicoherent search optimization using the KWS approximation for data from 2 detectors with duty cycle 70% and noise floor with fluctuations. The cost constraint is $C_0 = 472.0$ days. The best numerical solution is denoted with $\times$, with value: $\tilde{m}_{\text{opt}} \approx 0.05$, $\hat{m}_{\text{opt}} \approx 0.48$, $\Delta T_{\text{opt}} \approx 0.9$ days, $N_{\text{opt}} \approx 325.4$, $T_{\text{opt}} \approx 364.5$ days, $\tilde{n} = 2$, $\hat{n} = 3$, $h_{\text{th}}/\sqrt{S} \approx 4.4 \times 10^{-3} \sqrt{\text{Hz}}$. (a) KWS optimal total observation time T and number of segments N . (b) KWS optimal maximal mismatch on the coarse and fine grid, $\tilde{m}$ resp. $\hat{m}$.

5.3.2. Directed search using real data

In this subsection we apply the optimization procedure to real data collected by the Hanford (H1) and Livingston (L1) LIGO detectors during the S5 run. The most sensitive data is found around 169.875 Hz, thus the optimization will be done at this frequency. The details about the data are summarized in Table 5.1. It spans 653 days in 17797 SFTs of duration $T_{\text{SFT}} = 1800$ s. With this the average duty cycle is approximately 0.28 in each detector.

run	detector	f [Hz]	first SFT	last SFT	N_{SFT}	T_{span} [d]
S5	H1	169.875	818845553	875277921	9331	653
S5	L1	169.875	818845553	875278812	8466	653

Table 5.1.: Detector data used to test the numerical optimization under real conditions.

Case D: Keeping the cost constraint

We first keep the cost constraint equal to the computing cost used in the examples with simulated data, namely $C_0 \approx 472$ days. The result of the optimization procedure is plotted in Fig. 5.6. The minimal measurable strength is

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 5.1 \times 10^{-3} \sqrt{\text{Hz}} , \quad (5.46)$$

with the optimal search parameters

$$\tilde{m}_{\text{opt}} \approx 0.05 , \quad \hat{m}_{\text{opt}} \approx 0.49 , \quad (5.47)$$

and segment number and duration

$$\begin{aligned} N_{\text{opt}} &\approx 105 , \quad \Delta T_{\text{opt}} \approx 1.0 \text{ days} , \\ T_{\text{opt}} &\approx 498.2 \text{ days} . \end{aligned} \quad (5.48)$$

The sensitivity gain compared to the fully-coherent solution is ≈ 1.5 .

Case E: Using Einstein@Home

We now consider using the Einstein@Home distributed computing environment to increase the computing cost constraint to $C_0 = 360 \times 10^3$ days on a single computing core. Such computing power corresponds to approximately 30 days on 12000 24x7 single core CPUs. Given that in the WSG approximation the sensitivity scales with the computing cost like

$$h_{\text{th}}^{-1} \propto C_0^{1/10} \quad (5.49)$$

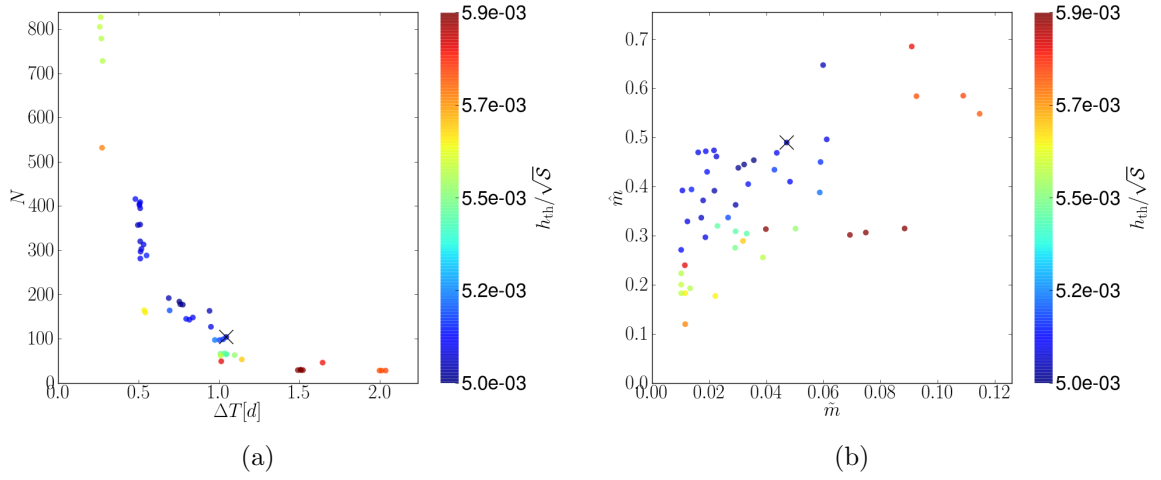


Figure 5.6.: Semicoherent search optimization using the KWS approximation with data from the H1 and L1 LIGO detectors during the S5 run at 169.875 Hz. The cost constraint is $C_0 = 472.0$ days. The best numerical solution is denoted with $\times$. The most sensitive optimal solution in the KWS approximation is: $\tilde{m}_{\text{opt}} \approx 0.05$, $\hat{m}_{\text{opt}} \approx 0.49$, $\Delta T_{\text{opt}} \approx 1.0$ days, $N_{\text{opt}} \approx 105.0$, $T_{\text{opt}} \approx 498.2$ days, $\tilde{n} = 2$, $\hat{n} = 3$, $h_{\text{th}}/\sqrt{S} \approx 5.1 \times 10^{-3} \sqrt{\text{Hz}}$. (a) KWS optimal total observation time T and number of segments N . (b) KWS optimal maximal mismatch on the coarse and fine grid, $\tilde{m}$ resp. $\hat{m}$.

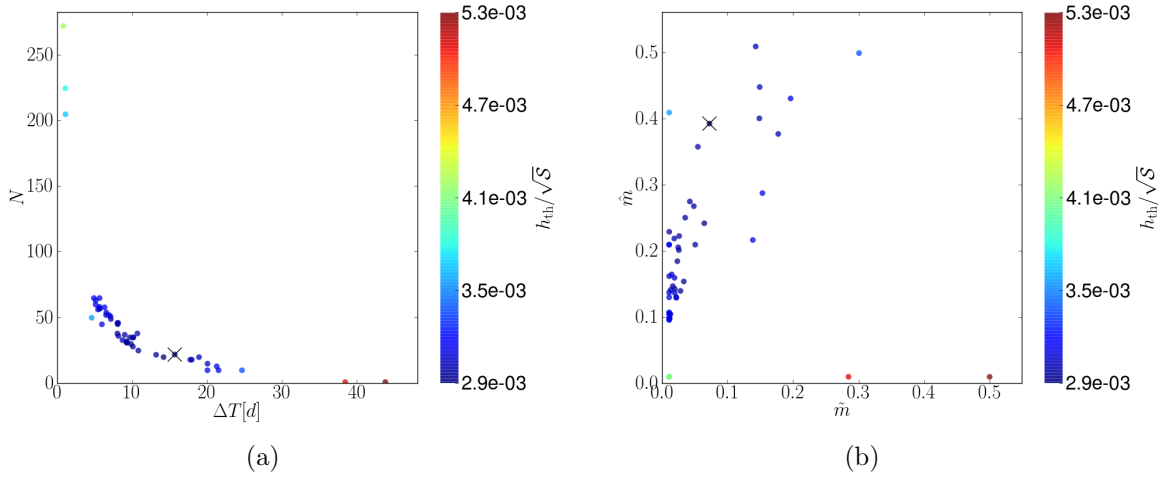


Figure 5.7.: Semicoherent search optimization using the KWS approximation with data from the H1 and L1 LIGO detectors during the S5 run at 169.875 Hz. The cost constraint is $C_0 = 360 \times 10^3$ days. The best numerical solution is denoted with $\times$. The most sensitive optimal solution in the KWS approximation is: $\tilde{m}_{\text{opt}} \approx 0.07$, $\hat{m}_{\text{opt}} \approx 0.39$, $\Delta T_{\text{opt}} \approx 15.6$ days, $N_{\text{opt}} \approx 22$, $T_{\text{opt}} \approx 494.6$ days, $\tilde{n} = 2$, $\hat{n} = 3$, $h_{\text{th}}/\sqrt{S} \approx 2.9 \times 10^{-3} \sqrt{\text{Hz}}$. (a) KWS optimal total observation time T and number of segments N . (b) KWS optimal maximal mismatch on the coarse and fine grid, $\tilde{m}$ resp. $\hat{m}$.

we expect to achieve sensitivity of $\approx 1.54 \times 10^{-3} \sqrt{\text{Hz}}$. In fact the most sensitive numerical solution using the WSG approximation yields $h_{\text{th}}^{\text{WSG}} \approx 1.6 \times 10^{-3} \sqrt{\text{Hz}}$.

The result of the numerical optimization procedure when using the KWS approximation is plotted in Fig. 5.7. The minimal measurable strength is

$$\left. \frac{h_{\text{th}}^{\text{KWS}}}{\sqrt{\mathcal{S}}} \right|_{\text{opt}} \approx 2.9 \times 10^{-3} \sqrt{\text{Hz}} , \quad (5.50)$$

with the optimal search parameters

$$\tilde{m}_{\text{opt}} \approx 0.07 , \quad \hat{m}_{\text{opt}} \approx 0.39 , \quad (5.51)$$

and segment number and duration

$$\begin{aligned} N_{\text{opt}} &\approx 22 , \quad \Delta T_{\text{opt}} \approx 15.6 \text{ days} , \\ T_{\text{opt}} &\approx 494.6 \text{ days} . \end{aligned} \quad (5.52)$$

The gain in sensitivity compared to the fully coherent solution is ≈ 2.64 . Note, that by the enormous increase of the computing power, we improve the gain in sensitivity compared to the case of ideal data. In a sense we “fill” the gaps in the data with computing cycles.

Case	T_{span} [days]	ϵ	data	C [days]	$\tilde{m}_{\text{opt}}$	$\hat{m}_{\text{opt}}$	ΔT_{opt} [days]	N_{opt}	$h_{\text{th}}^{\text{KWS}}/\sqrt{\mathcal{S}}$ [$\sqrt{\text{Hz}}$]	gain
A	365	1	const	472	0.10	0.29	1.3	188	3.5×10^{-3}	2.2
B	365	0.7	const	472	0.04	0.46	0.8	364	4.4×10^{-3}	1.74
C	365	0.7	var	472	0.04	0.45	0.8	366	4.4×10^{-3}	1.74
D	653	0.28	S5	472	0.05	0.49	1.0	105	5.1×10^{-3}	1.5
E	653	0.28	S5	360×10^3	0.07	0.39	15.6	22	2.9×10^{-3}	2.64

Table 5.2.: Summary of the results of the optimization. T_{span} is the time spanned by the data set used to determine the optimal search parameters. The last column gives the gain in sensitivity compared to the fully coherent solution, Eq. (5.30).

5.4. Discussion

In this chapter we studied the optimization of semicoherent searches for continuous gravitational waves, in particular the StackSlide search, at constrained computing cost under real conditions, by taking into account possible gaps in the data and noise level changes. The presented method to obtain optimal search parameters is based on numerical optimization, in this particular case a MADS algorithm. The outcome of the optimization procedure is the set of the optimal search parameters $\{\tilde{m}, \hat{m}, N, \Delta T\}$ as well as the selected data and an estimate of the expected sensitivity.

We showed that under ideal conditions, our numerical optimization method recovers the optimal solution found by using the analytical method discussed in Chap. 4. The

optimization procedure is immediately applicable to searches over simple (nearly) "box" parameter-space shape. While the proposed optimization method can be easily adapted to other types of searches, by modification of the computing cost function, further work is required to extend the applicability of the optimization procedure to an arbitrary parameter-space shape. The proposed optimization method assumed a fixed frequency band. Further work is required to relax this condition, in order to answer the question, what is the optimal trade-off between the size of the searched parameter space (width of the search) and its coverage (depth of the search).

Part III.

Fully coherent Follow-up Studies

Summary

In order to confirm or discard a candidates selected in a semicoherent wide parameter-space search as a real gravitational-wave event, the candidate requires further fully coherent follow-up investigation.

In Chap. 6 we derive two different methods to compute the minimal required integration time of a fully coherent follow-up of candidates produced in wide parameter space semicoherent searches. We numerically compare these methods in terms of integration duration and computing cost. In a Monte Carlo study we confirm that we can achieve the required detection probability.

In Chap. 7 we present a general method and implementation of a *direct* (two-stage) transition to a fully coherent follow-up of semicoherent candidates using *all* of the available data. This method is based on a grid-less Mesh Adaptive Direct Search algorithm using the $\mathcal{F}$ -statistic. We demonstrate the detection power and computing cost of this follow-up procedure using extensive Monte Carlo simulations on (simulated) semicoherent candidates from a directed as well as from an all-sky search setup.

In Chap. 8 we report on a fully coherent follow-up of some loud candidates from the Einstein@Home all-sky search for periodic gravitational waves in LIGO fifth Science Run (S5) data [59] using the follow-up method developed in Chap. 7. We show the efficiency of the follow-up pipeline using simulated signals added into simulated Gaussian noise data. We discuss the impact of possible deviation of the second-order spindown from zero on the zoom stage, given that the zoom stage is restricted only to the first-order spindown, but the observation time is long, approx. 264 days. We recover three simulated signals added at the hardware level into the gravitational wave detectors and identified by the Einstein@Home search. Finally, we follow-up a set of nine loud candidates not associated with simulated signals. No significant gravitational-wave events have been found.

6. Minimal required observation time

Isolated neutron stars as potential sources of continuous gravitational waves are optimally studied with fully coherent matched filtering methods. These methods are not directly applicable to previously unknown objects due to the large parameter space that needs to be covered in all-sky wide parameter-space searches and the related enormous computing cost [53]. Advanced semicoherent techniques, e.g., StackSlide searches on the distributed computing environment Einstein@Home [102], produce candidates that require follow-up in greatly reduced parameter space regions. A follow-up scheme consists of two basic stages. In the first *refinement* stage, we find the maximum-likelihood estimator and associated optimal search volume V_0 . In the second *zoom* stage, we zoom in on the optimal search volume by semicoherent or fully coherent integration. In this chapter we focus on a fully coherent zoom for which we derive and discuss two different methods to compute the minimal required coherent integration time in order to distinguish real signals from noise.

6.1. Properties of $\mathcal{F}$ -statistic searches

The $\mathcal{F}$ -statistic was first derived in [28] for the single detector case and generalized to multi-detector searches in [46]. Continuous gravitational-wave signals are monochromatic and sinusoidal in the frame of the gravitational-wave source and undergo phase- and amplitude modulation due to the rotation and orbital motion of the detector. The $\mathcal{F}$ -statistic is analytically amplitude-maximized, thus the parameter space to search for signals is spanned by the remaining “Doppler parameters” λ , namely sky position (α - right ascension, δ - declination) and intrinsic frequency and frequency derivatives ($f, \dot{f}, \ddot{f} \dots$), further referred to as spindowns. Searching for previously unknown objects with matched filtering implies computing matched filters for different points in parameter space, also referred to as templates. As realized in [52, 51] in the context of searches for gravitational waves from inspiraling binaries, a geometrical approach is best suited for optimal template placement and template counting. This is made possible by the introduction of a metric tensor g_{ij} on the parameter space and mismatch μ

$$\mu = g_{ij} \Delta \lambda^i \Delta \lambda^j + \mathcal{O}(\Delta \lambda^3) , \quad (6.1)$$

where the mismatch μ measures the fractional loss of (squared) signal to noise ratio (SNR) ρ^2 due to the usage of a nearby template λ_c with offset $\Delta \lambda = \lambda_c - \lambda_s$ from the

true parameters of a putative signal λ_s

$$\mu = \frac{\rho_s^2 - \rho_c^2}{\rho_s^2} , \quad (6.2)$$

with the squared SNR ρ_s^2 and ρ_c^2 obtained at point λ_s and λ_c , respectively. Given the metric, the problem of efficient lattice and alternative random and stochastic template-bank construction is studied in [54, 55, 103].

6.1.1. Fully coherent search

A fully coherent search is the classical and most sensitive $\mathcal{F}$ -statistic-based search in the case of unlimited available computing power or a sufficiently cheap computing cost requirement. The squared SNR ρ^2 scales linearly with the observation time T , according to the following formula:

$$\rho^2 = h_0^2 R N_d T S^{-1}(f) , \quad (6.3)$$

where h_0 is the intrinsic signal amplitude, R represents the geometrical “detector response” , S is the one-sided noise spectral density, which is assumed constant in a narrow frequency band around f , and N_d is the number of detectors [44]. In the presence of a signal, the $\mathcal{F}$ -statistic follows a non-central χ^2 distribution with four degrees of freedom and non-centrality parameter ρ^2 . Thus the expectation value is

$$E[2\mathcal{F}_s] = 4 + \rho^2 , \quad (6.4)$$

with standard deviation

$$\sigma[2\mathcal{F}_s] = \sqrt{2(4 + 2\rho^2)} . \quad (6.5)$$

6.1.2. Semicoherent search

At fixed and limited computing cost a more sensitive detection statistic can be constructed from the incoherent combination of results obtained by coherent integration of shorter data segments. In particular we consider a Stack-Slide search [64, 88, 98], where the statistic is the sum of the $\mathcal{F}$ -statistic over the segments:

$$\Sigma = \sum_{k=1}^N 2\mathcal{F}_k(\lambda) . \quad (6.6)$$

This new statistic Σ follows a non-central χ^2 distribution with $4N$ degrees of freedom, thus the expectation value is

$$E[\Sigma] = 4N + \rho_\Sigma^2 , \quad (6.7)$$

where the non-centrality parameter is the sum of the squared SNRs over different segments

$$\rho_\Sigma^2 = \sum_{k=1}^N \rho_k^2 . \quad (6.8)$$

A trivial but useful reformulation of Eq. (6.7) is in terms of average $\overline{2\mathcal{F}} = \frac{1}{N} \sum_k 2\mathcal{F}_k$ and $\overline{\rho^2} = \frac{1}{N} \sum_k \rho_k^2$, namely

$$E[\overline{2\mathcal{F}}] = 4 + \overline{\rho^2} . \quad (6.9)$$

6.1.3. Template counting

The number of templates sufficient to cover the search volume V_0 is given by [54]

$$\mathcal{N}_n = \theta m^{-n/2} \mathcal{V}_n , \quad (6.10)$$

where θ is the normalized thickness characterizing the geometric structure of covering, m is the maximum allowed mismatch, n the number of dimensions and

$$\mathcal{V}_n = \int d^n \lambda \sqrt{\det g} , \quad (6.11)$$

is the metric template-bank volume with g_{ij} the parameter space metric. This is the general form of the template counting formula, which is valid for arbitrary lattices and curved parameter spaces. In practice, using the flat metric approximation, where the metric coefficients are constant, we can take the determinant out of the integral. Moreover, if the parameter space is a n -dimensional “box”, we can replace the integral over infinitesimal displacement $d\lambda$ by a product of n “search bands” $\Delta\lambda$, namely

$$\mathcal{V}_n = \sqrt{\det g} \prod_{i=1}^n \Delta\lambda_i . \quad (6.12)$$

Follow-up of candidates from semicoherent searches involves a semicoherent metric, shown in [64, 56] to be the average of the metric computed for every segment. The semicoherent metric allows us to estimate the search band $\Delta\lambda^i$ around the follow-up candidate using the diagonal elements of the inverse Fisher matrix [41, 104], i.e.,

$$\Delta\lambda^i \equiv \kappa \sqrt{\{\Gamma^{-1}\}^{ii}} , \quad (6.13)$$

with

$$\{\Gamma^{-1}\}^{ii} = g^{ii} / \rho^2 , \quad (6.14)$$

where κ defines the confidence level and g^{ij} is the inverse matrix to g_{ij} ¹. In this chapter we use an analytical semicoherent metric first derived by Pletsch [56]. For coherent integration time longer than a day, but much shorter than a year, using sky coordinates proportional to the light travel time from the Earth’s center to the detector $\tau_E \approx 21 \times 10^{-3} s$, the number of sky templates at fixed frequency f converges to [56]

$$\mathcal{N}_{sky} = \frac{2\pi^3 \tau_E^2 f^2}{m} . \quad (6.15)$$

¹In this chapter we compute the Fisher matrix using the semicoherent metric rescaled by $\overline{\rho^2}$ instead of ρ_Σ^2 , which can be interpreted as a conservative choice; see Chap. 7.

The semicoherent parameter space is finer than the coherent one by a refinement factor γ . Using the notion of *refinement per direction* γ_n we can also obtain the search bands from the extents of the fully coherent metric, namely

$$\Delta\lambda^i = \kappa \sqrt{\frac{g^{ii}}{\gamma_i^2 \rho^2}} . \quad (6.16)$$

For uniformly distributed segments of data without gaps, based on [56] the refinement factors can be obtained as

$$\gamma_f = 1 , \quad (6.17)$$

$$\gamma_{\dot{f}} = \sqrt{5N^2 - 4} , \quad (6.18)$$

$$\gamma_{\ddot{f}} = \sqrt{(35N^4 - 140N^2 + 108)/3} , \quad (6.19)$$

$$\gamma_{\dot{f}\ddot{f}} = \sqrt{(105N^8 - 1260N^6 + 5012N^4 - 6160N^2 + 2304)/(5N^2 - 4)} . \quad (6.20)$$

Finally, for simplicity of the template-bank construction, we use a hyper-cubic lattice to place templates, though hyper-cubic lattices are in general suboptimal, compared to better solutions, e.g., A_n^* lattice. The normalized thickness for an n -dimensional hyper-cubic grid is [54]

$$\theta_n = n^{n/2} 2^{-n} . \quad (6.21)$$

The proper choice of the number of dimensions that maximizes the number of templates [53, 64, 88, 98] $\mathcal{N}$ is:

$$\mathcal{N} = \max_n \mathcal{N}_n . \quad (6.22)$$

6.1.4. Computing cost

In the follow-up of real candidates, especially weak signal candidates, along with the constraint of the total amount of available data, the computing cost constraint may limit significantly the feasibility of the search. Thus the computing-cost requirement is of particular interest. There are currently two different strategies to implement an $\mathcal{F}$ -statistic search code in LIGO's reference software suite `lalsuite`[105], namely the SFT-method based on short Fourier transforms of the data with duration T_{SFT} [44] and the FFT-method based on barycentric resampling [97]. Regarding the computational cost, the FFT method is preferable, as the computational requirement to calculate the $\mathcal{F}$ -statistic, for a single point in the parameter space, scales only with $\log T$, while the cost of the SFT algorithm scales with T . However, for historical reasons the SFT method is currently still more often used by LIGO/LSC [106, 79, 63], is well tested and we can use recent timing information. The computing cost of a SFT-based $\mathcal{F}$ -statistic search is

$$C = \mathcal{N}_{c_0} N_{\text{SFT}} , \quad (6.23)$$

where N_{SFT} is the number of used SFTs, namely

$$N_{\text{SFT}} = N_d T / T_{\text{SFT}} \quad (6.24)$$

and c_0 is the fundamental implementation- and hardware-specific computing constant per SFT and template.

6.2. Minimal required observation time

The main scope of the present work is to find the minimal required observation time that guarantees a certain detection probability of a putative signal buried deep in the detector noise at a certain confidence level by using the fully coherent $\mathcal{F}$ -statistic search technique. We consider two different methods to compute the required integration duration. In method 1, which is closely related to hypothesis testing, we use the concept of false-alarm and false-dismissal probability to achieve certain detection probability. This is the natural way to compute the required integration time. In method 2 we alternatively use the more intuitive notion of expectation value to find the observation duration that guarantees the required detection probability.

6.2.1. Method 1

In absence of a signal, the probability density function of the $\mathcal{F}$ -statistic reduces to a central χ^2 -distribution, and the false-alarm probability is given by

$$p_{\text{fA}}^1 = \int_{2\mathcal{F}_{\text{th}}}^{\infty} d(2\mathcal{F}) \chi_4^2(2\mathcal{F}; 0) , \quad (6.25)$$

where p_{fA}^1 denotes single trial false-alarm probability and $\chi_4^2(2\mathcal{F}; 0)$ is the central χ^2 -distribution with 4 degrees of freedom. The integration of $\chi_4^2(2\mathcal{F}, 0) = \frac{1}{2}\mathcal{F}e^{-\mathcal{F}}$ yields

$$p_{\text{fA}}^1 = (1 + \mathcal{F}_{\text{th}})e^{-\mathcal{F}_{\text{th}}} . \quad (6.26)$$

The overall false-alarm probability of crossing the threshold $2\mathcal{F}_{\text{th}}$ in $\mathcal{N}$ trials is

$$p_{\text{fA}} = 1 - (1 - p_{\text{fA}}^1)^{\mathcal{N}} \approx p_{\text{fA}}^1 \mathcal{N} , \quad (6.27)$$

when $p_{\text{fA}}^1 \mathcal{N} \ll 1$ [28, 107], thus

$$p_{\text{fA}}^1 = p_{\text{fA}} / \mathcal{N} . \quad (6.28)$$

We cannot solve Eq. (6.26) analytically, but numerical solution gives a threshold $2\mathcal{F}_{\text{th}}$ value. This allows us to numerically integrate the false-dismissal probability

$$p_{\text{fD}}(2\mathcal{F}_{\text{th}}, \rho^2) = \int_{-\infty}^{2\mathcal{F}_{\text{th}}} (d2\mathcal{F}) \chi_4^2(2\mathcal{F}, \rho^2) , \quad (6.29)$$

where $p_{\text{fD}}(2\mathcal{F}_{\text{th}}) = 1 - p_{\text{det}}$, with the desired detection probability p_{det} and $\chi_4^2(2\mathcal{F}, \rho^2)$ is the non-central χ^2 -distribution with 4 degrees of freedom and non-centrality parameter

ρ^2 . At fixed p_{fA}^* and p_{fD}^* , using the above equation, we can compute a threshold SNR $\rho_{\text{th}}(p_{\text{fA}}^*, p_{\text{fD}}^*)$. The required T is such that the inequality

$$\rho_{\text{ac}}^2(T) \geq \rho_{\text{th}}^2(p_{\text{fA}}^*, p_{\text{fD}}^*) \quad (6.30)$$

holds, where $\rho_{\text{ac}}^2(T)$ is the accumulated SNR due to the presence of signal in the analyzed data. Assuming that the follow-up search will use data of similar constant noise floor, we can rewrite Eq. (6.3) as

$$\rho_{\text{ac}}^2(T) = \rho_c^2 \frac{N_d T}{N_d^c \Delta T} , \quad (6.31)$$

where ΔT is the length of one segment in the semicoherent search using data from N_d^c number of detectors. With the average $\overline{2\mathcal{F}_c}$ value of the candidate, we can compute its SNR ρ_c from Eq. (6.9), namely

$$\rho_c^2 = E[\overline{2\mathcal{F}_c}] - 4 . \quad (6.32)$$

Substitution in the equations above yields the accumulated SNR in presence of signal

$$\rho_{\text{ac}}^2 = (E[\overline{2\mathcal{F}_c}] - 4) \frac{N_d T}{N_d^c \Delta T} , \quad (6.33)$$

which gives the required minimal T .

6.2.2. Method 2

Computation of the $\mathcal{F}$ -statistic on data with no signal, has a certain expectation value, therefore we ask what is the expected maximal $2\mathcal{F}$ value $E[2\mathcal{F}_{\mathcal{N}}]$ in $\mathcal{N}$ trials in Gaussian noise, where $\mathcal{F}_{\mathcal{N}} \equiv \max \{\mathcal{F}\}_{i=1}^{\mathcal{N}}$. The probability to get $(\mathcal{N} - 1)$ values of $2\mathcal{F}$ less than $2\mathcal{F}_{\mathcal{N}}$ follows a binomial distribution, namely

$$p_{\mathcal{N}}(2\mathcal{F}_{\mathcal{N}}) = \binom{\mathcal{N}}{1} \chi_4^2(2\mathcal{F}, 0) (1 - \alpha_1)^{\mathcal{N}-1} \quad (6.34)$$

$$= \frac{1}{2} \mathcal{N} \mathcal{F}_{\mathcal{N}} e^{-\mathcal{F}_{\mathcal{N}}} (1 - (1 + \mathcal{F}_{\mathcal{N}}) e^{-\mathcal{F}_{\mathcal{N}}})^{\mathcal{N}-1} . \quad (6.35)$$

With this we can numerically integrate the expectation value

$$E[2\mathcal{F}_{\mathcal{N}}] = \int_0^\infty d(2\mathcal{F}_{\mathcal{N}}) 2\mathcal{F}_{\mathcal{N}} p_{\mathcal{N}}(2\mathcal{F}_{\mathcal{N}}) , \quad (6.36)$$

and standard deviation

$$\sigma_{\mathcal{N}}(2\mathcal{F}_{\mathcal{N}}) = \left(\int_0^\infty d(2\mathcal{F}_{\mathcal{N}}) (2\mathcal{F}_{\mathcal{N}} - E[2\mathcal{F}_{\mathcal{N}}])^2 p_{\mathcal{N}}(2\mathcal{F}_{\mathcal{N}}) \right)^{1/2} . \quad (6.37)$$

To safely distinguish a real signal from pure noise, we can require the following inequality to hold:

$$E[2\mathcal{F}_s] - h\sigma_s(2\mathcal{F}_s) > E[2\mathcal{F}_{\mathcal{N}}] + h\sigma_{\mathcal{N}}(2\mathcal{F}_{\mathcal{N}}) , \quad (6.38)$$

where the expectation value $E[2\mathcal{F}_s]$ of a real signal and its standard deviation $\sigma_s(2\mathcal{F}_s)$ are computed using Eqs. (6.4) and (6.5). As all terms in inequality (6.38) are function of the observation time, this gives an alternative method to compute the minimal required integration time. Fine-tuning of Eq. (6.38) is possible through the safety parameter h , which we quantify by using Chebyshev's inequality. For a random variable X , with expected value $E[X]$ and standard deviation σ ,

$$P(|X - E[X]| \geq h\sigma) \leq 1/h^2, \quad (6.39)$$

which means that at least a fraction

$$p = 1 - 1/h^2 \quad (6.40)$$

of the data is within h standard deviations on either side of the mean [108]. Rearranging the above equation yields

$$h = 1/\sqrt{1-p}. \quad (6.41)$$

Having two independent random variables, $2\mathcal{F}_s$ and $2\mathcal{F}_N$, we can label the fraction of data around each mean as p_s and p_N and introduce the joint probability $p_J = p_s p_N$. We see, that the same joint probability can be achieved for different combinations of p_s and p_N . However, a natural choice is $p_s = p_N$, thus

$$h = 1/\sqrt{1 - \sqrt{p_J}}. \quad (6.42)$$

We give a set of p_J values and related h in Table 6.1.

p_J	0.75	0.90	0.95	0.99
h	2.73	4.41	6.28	14.12

Table 6.1.: Joint probability p_J and corresponding required h standard deviations.

Fixing p_J to some value and with this h in inequality [(6.38)], we can compute the minimal required coherent observation time T , such that [(6.38)] holds. For this integration time, the joint probability p_J becomes the separation probability $p_{\text{sep}} = p_J$. This is the probability, that a candidate due to the presence of a signal is consistent with the signal hypothesis and a candidate due to the noise is consistent with the noise hypothesis. Taking into account that $p_s = 1 - p_{\text{fD}}$ and $p_N = 1 - p_{\text{fA}}$, we find the relation of the separation probability to the detection probability, namely $p_{\text{sep}} = p_{\text{det}}(1 - p_{\text{fA}})$, or for negligible false-alarm $p_{\text{det}} \approx p_{\text{sep}}$.

6.3. Method comparison

6.3.1. Numerical predictions

In the following we compare the two methods to find the minimal required integration time described in the previous section in terms of observation duration and computing

cost. We consider a StackSlide search with $N = 205$ segments of duration $\Delta T = 25$ hours, each using data from $N_d^c = 2$ detectors. For a hypothetical candidate with fixed Doppler parameters $\alpha = 1.45$ rad, $\delta = 0$ rad $f = 185$ Hz, $\dot{f} = -1 \times 10^{-9}$ Hz/s, we pick an average strength in the range $\overline{2\mathcal{F}_c} \in [5, 13]$. Then using Eq. (6.13) with $\kappa = 1$ and the semicoherent metric we compute the search bands associated with such a candidate. Having that, for mismatch $m = 0.01$ and a hyper-cubic lattice, we can compute the number of templates using Eq. (6.10) and the fully coherent metric. Using method 1, requiring detection probability $p_{\text{det}}^* = 0.9$ at overall false-alarm probability $p_{\text{fA}}^* = 0.01$ using Eq. (6.29) we compute $\rho_{\text{th}}^2(p_{\text{fA}}^*, p_{\text{fD}}^*)$ and the minimal required observation time T_1 , which substituted in Eq. (6.33) with $N_d = N_d^c$ satisfies Eq. (6.30). For method 2 a separation probability equal to p_{det}^* yields safety factor $h = 4.41$, see Table 6.1. We label the integration time that satisfies Eq. (6.38) as T_2 and plot both integration times $T_1(\overline{2\mathcal{F}_c})$ and $T_2(\overline{2\mathcal{F}_c})$ in Fig. 6.1 (a) as function of $\overline{2\mathcal{F}_c}$. With the number of templates for T_1 and T_2 we estimate the computing cost C_1 and C_2 using the fundamental computing cost constant $c_0 = 7 \times 10^{-8}$ s in Eq. (6.24) and assuming SFTs of duration $T_{\text{SFT}} = 1800$ s in Eq. (6.25). $C_1(\overline{2\mathcal{F}_c})$ and $C_2(\overline{2\mathcal{F}_c})$ are plotted in Fig. 6.1 (b). In Fig. 6.1 (c) we plot how the expectation value from a real signal grows with increasing T compared to loudest candidate from Gaussian noise. In this plot the candidate strength is fixed to $\overline{2\mathcal{F}_c} = 8.5$. We see that method 2 yields much longer observation time, at same candidate strength compared to method 1. Due to the resulting much larger number of templates, the computing cost, especially for weak candidates, is much higher. The inferiority of method 2 compared to method 1 in terms of required integration duration and computing power can be explained by the *ad hoc* construction of method 2 and the use of Chebyshev's inequality, which is only a lower bound. In this sense method 2 is a more conservative approach, though the important information about false-alarm and false-dismissal probability gets lost in this framework. The computing cost of method 1 looks very promising even for weak candidates, however we should keep in mind that this is lower limit and the cost of a search with real data would most likely be much higher. The reason for this is that gaps in the data are direct penalty for the growth of ρ_{ac}^2 , while ρ_{th}^2 remains unaffected. Furthermore, for very weak signals, the required integration duration may violate the assumption of constant sky resolution, thus we would underestimate the number of templates, resulting in a higher false-dismissal.

6.3.2. Monte Carlo results

To confirm the numerical predictions of method 1 we perform the following Monte Carlo studies. We create a set of 205 segments with duration 25 hours of Gaussian noise and draw a set of pulsar parameters $\alpha \in (0, 2\pi)$, $\delta \in (-\pi/2, \pi/2)$, $\cos \iota \in (-1, 1)$, $\psi \in (0, 2\pi)$, $\phi_0 \in (0, 2\pi)$ at fixed frequency of $f = 185$ Hz and spindown value in the range $\dot{f} \in (-f/\tau, 0)$, where $\tau = 2220$ y is the minimal spindown age of the source [53]. We inject a signal with the above parameters and intrinsic signal amplitude h_0 high enough to produce a candidate with expected average strength $E[2\mathcal{F}_s] \in [12, 13]$. To find the actual injected value we first do a targeted StackSlide search at the point of the

6. Minimal required observation time

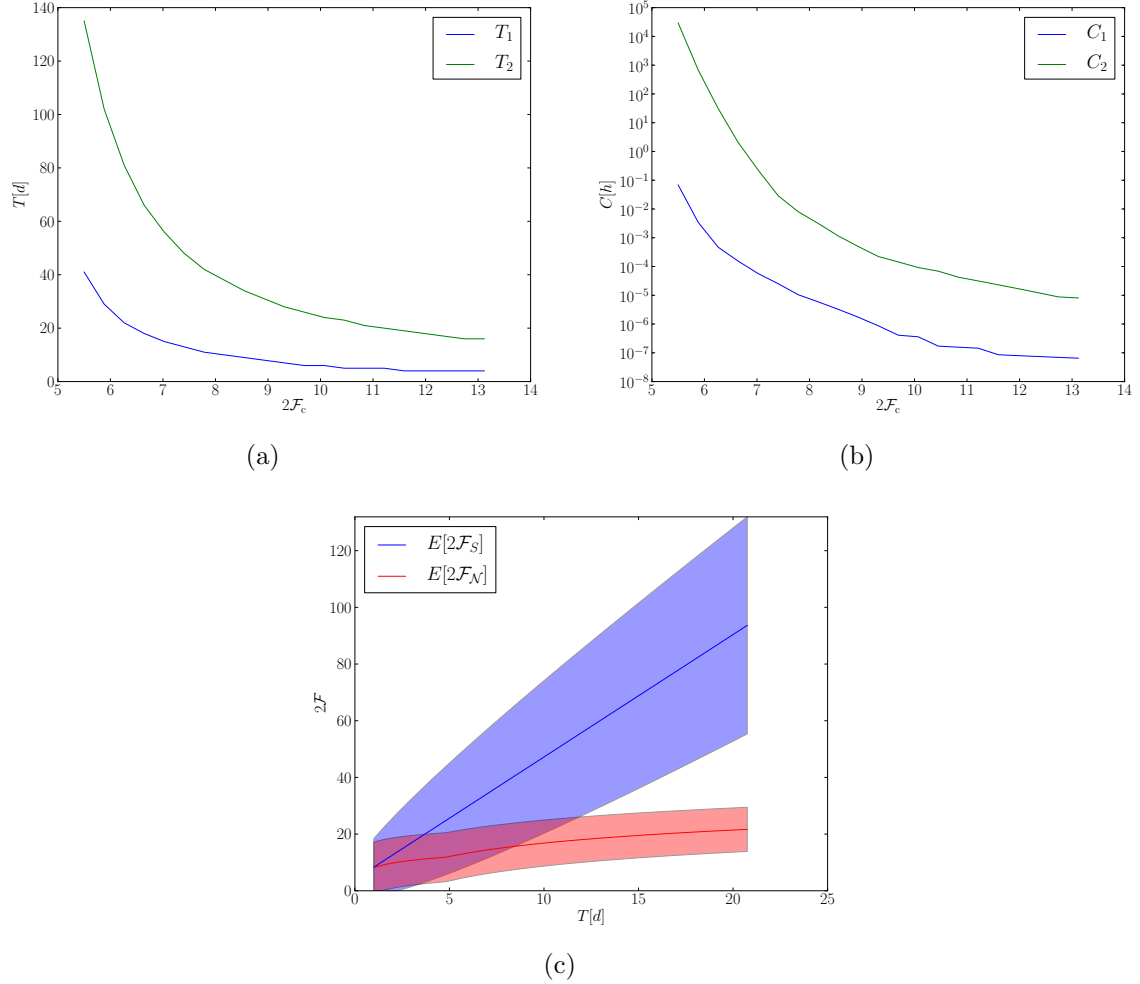


Figure 6.1.: Numerical comparison between method 1 and method 2 (quantities labeled with 1 and 2, respectively). (a) Required coherent integration time as function of the strength of the candidate. (b) Computing cost depending on the strength of the candidate. (c) Expected value of signal, noise and related $h = 4.41$ standard deviations for detection probability $p_{\text{det}} = 0.9$ of a candidate with $2\overline{\mathcal{F}}_c = 8.5$.

injection. With this measured injected $\overline{2\mathcal{F}}_s$ value, using Eq. (6.13) we compute Fisher extents, from which we draw a random parameter point λ_c satisfying

$$\Gamma_{ij}\Delta\lambda^i\Delta\lambda^j < 1 . \quad (6.43)$$

The point λ_c is within the $1\text{-}\sigma$ Fisher ellipsoid of the true signal location and becomes the candidate to follow up. Following the scheme for method 1 as described above, we compute the minimal required coherent observation time targeting detection probability $p_{\text{det}}^* = 0.9$ and search for the signal. After computation of $2\mathcal{F}_s$ using the data with the injected signal, we compute $2\mathcal{F}_N$ with the same grid and integration duration using the noise only data. We claim “detection” whenever the loudest measured $2\mathcal{F}_s$ value in the data with injected signal is higher than the loudest measured $2\mathcal{F}_N$ of the noise. The result of the Monte Carlo simulations is as follows: in 897, out of 1000 trials, the measured $2\mathcal{F}_s$ value in the data containing injected signal exceeds the measured $2\mathcal{F}_N$ value of the noise only data. With this the achieved detection probability $p_{\text{det}} = 0.897 \pm 0.023$ is in accordance with the targeted detection probability p_{det}^* .

6.4. Discussion

We derived two different methods to compute the minimal required coherent integration time in a fully coherent $\mathcal{F}$ -statistic search in the *zoom* stage of follow-up of candidates from a semicoherent StackSlide search. By numerical comparison we showed that method 1 is superior to method 2 in terms of required integration duration and computing cost. We confirmed in a Monte Carlo study that the predicted coherent integration time is sufficient to achieve the desired detection probability. The results of this chapter have been derived for Gaussian data without gaps and two detectors of equal noise floor. Further extension of this work would require the usage of data selection procedure, however this complication can be worked around, e.g., by using all the data, as proposed in the next chapter.

7. Fully coherent follow-up of continuous gravitational-wave candidates

Continuous gravitational waves (CWs) are expected to be emitted from rapidly spinning nonaxisymmetric compact objects, e.g., neutron stars. The computational cost of a coherent matched-filtering detection statistic, such as the $\mathcal{F}$ -statistic [28], is small provided the parameters of the source (i.e., sky position α, δ , frequency f , frequency derivatives $\dot{f}, \dots$) are known. However, wide parameter-space searches for unknown sources quickly become computationally prohibitive, due to the large number of points in parameter space (templates) that need to be searched [53].

In order to first reduce the parameter space to smaller, more promising regions, semicoherent search techniques have been developed [64, 88, 43, 92] and are currently being used [109, 59], for example in the Einstein@Home distributed computing environment [102]. In a semicoherent search, the total amount of data T is divided into N shorter segments of duration ΔT . The coherent statistics from the individual segments are combined to a new semicoherent statistic. At a fixed computing cost these semicoherent methods are (typically) more sensitive than fully coherent searches [98].

Structuring a wide parameter-space search into hierarchical stages, which increasingly concentrate computational power onto the more promising regions of parameter space, was first described in [53] and elaborated further in [64], where a two-stage semicoherent hierarchical search was considered. An extended hierarchical scheme with an arbitrary number of semicoherent stages and a final fully coherent stage was studied numerically in [88], which concluded that three semicoherent stages will typically be a good choice. In [110] and [111] the use of an optimization procedure has been considered in the process of estimation of the source parameters, once a candidate is considered as a detection. In both cases, however, no practical method or implementation was provided for the systematic coherent follow-up of semicoherent candidates.

The aim of the present chapter is to introduce such a coherent follow-up search strategy and implementation. This is achieved by exploring the parameter space around a semicoherent candidate using a Mesh Adaptive Direct Search (MADS) algorithm. Using this method, we find that a fully coherent follow-up (using all of the available data) of initial semicoherent candidates is computationally feasible.

This chapter is organized as follows: in Sec. 7.1 we describe the relevant basic concepts in CW searches, in Sec. 7.2 we propose a search strategy for the systematic follow-up of

CW candidates, in Sec. 7.3 we present a Monte-Carlo study and in Sec. 7.4 we discuss the results.

Notation

We distinguish a quantity Q when referring to a fully coherent stage using a tilde, $\tilde{Q}$ and when referring to a semicoherent stage using an overhat, $\hat{Q}$. Averaging over segments is denoted by an overbar, $\overline{Q}$.

7.1. Continuous Gravitational Waves

Continuous gravitational-wave signals are quasi-monochromatic and sinusoidal in the source frame and undergo phase and amplitude modulation due to the diurnal and orbital motion of the detectors. The phase evolution of the signal at a detector can be approximated as [28]

$$\begin{aligned} \phi(t) \approx & \phi_0 + 2\pi \sum_{k=0}^s \frac{f^{(k)}(t_0)(t-t_0)^{k+1}}{(k+1)!} \\ & + 2\pi \frac{\vec{r}(t)}{c} \vec{n} \sum_{k=0}^s \frac{f^{(k)}(t_0)(t-t_0)^k}{k!}, \end{aligned} \quad (7.1)$$

where ϕ_0 is the initial phase, $f^{(k)} \equiv \frac{d^k f}{dt^k}$ are the derivatives of the signal frequency f at the solar system barycenter (SSB) at reference time t_0 , c is the speed of light, $\vec{r}(t)$ is the vector pointing from the SSB to the detector and $\vec{n}$ is the unit vector pointing from the SSB to the gravitational-wave source.

7.1.1. Detection statistic

Following [28, 46], the gravitational-wave response of a detector can be expressed as a sum over four (detector-independent) amplitude parameters multiplying four (detector-dependent) basis waveforms. The amplitude parameters can be analytically maximized over, and the resulting detection statistic, known as the $\mathcal{F}$ -statistic, is therefore a function only of the template “phase parameters” $\lambda \equiv \{\alpha, \delta, f, \dot{f}, \dots\}$, where α (right ascension) and δ (declination) denote the sky position of the source.

In the presence of a signal the fully coherent detection statistic $2\mathcal{F}$ follows a non-central χ^2 distribution with 4 degrees of freedom and a non-centrality parameter given by the squared signal-to-noise ratio (SNR), ρ^2 . The expectation value is therefore

$$E[2\mathcal{F}] = 4 + \rho^2, \quad (7.2)$$

with variance

$$\sigma^2[2\mathcal{F}] = 2(4 + 2\rho^2). \quad (7.3)$$

On the other hand, in the semicoherent approach, we divide the available data into N segments of duration ΔT and combine the individual coherent statistics of the segments to compute a semicoherent statistic, namely

$$\overline{2\mathcal{F}}(\lambda) = \frac{1}{N} \sum_{k=1}^N 2\mathcal{F}_k(\lambda) , \quad (7.4)$$

where $2\mathcal{F}_k$ is the coherent $\mathcal{F}$ -statistic in segment k . The quantity $N\overline{2\mathcal{F}}$ follows a non-central χ^2 distribution with $4N$ degrees of freedom, thus the expectation value of $\overline{2\mathcal{F}}$ is

$$E[\overline{2\mathcal{F}}] = 4 + \overline{\rho^2} , \quad (7.5)$$

with variance

$$\sigma^2[\overline{2\mathcal{F}}] = \frac{2}{N} (4 + 2\overline{\rho^2}) , \quad (7.6)$$

where $\overline{\rho^2}$ is the average SNR^2 over all segments, i.e.,

$$\overline{\rho^2} = \frac{1}{N} \sum_{k=1}^N \rho_k^2 , \quad (7.7)$$

and ρ_k^2 denotes the SNR^2 in segment k .

7.1.2. Mismatch and Fisher matrix

A search for sources with unknown signal parameters implies a loss of detection power compared to the perfectly matched case. To quantify this we use the notion of mismatch μ , as first introduced in [51, 52]. This is defined as the fractional loss of expected SNR^2 at some parameter space point λ compared to the expectation $\rho^2(\lambda_s)$ at the signal location λ_s , namely

$$\mu \equiv \frac{\rho^2(\lambda_s) - \rho^2(\lambda)}{\rho^2(\lambda_s)} , \quad (7.8)$$

such that $\mu \in [0, 1]$. Taylor expansion in small offsets $\Delta\lambda = \lambda - \lambda_s$ around the signal location yields

$$\mu \equiv g_{ij}(\lambda_s) \Delta\lambda^i \Delta\lambda^j + \mathcal{O}(\Delta\lambda^3) , \quad (7.9)$$

where implicit summation over repeated parameter-space indices i, j applies, and the symmetric positive-definite matrix g_{ij} is commonly referred to as the parameter-space *metric*.

Neglecting higher-order terms, one often uses the “metric mismatch approximation”, namely

$$\mu^* \equiv g_{ij}(\lambda_s) \Delta\lambda^i \Delta\lambda^j , \quad (7.10)$$

as a distance measure, with a range $\mu^* \in [0, \infty)$. This metric mismatch μ^* plays an important role in grid-based searches, where one typically constructs template banks

in such a way that the mismatch of any putative signal and the “closest” template is bounded by a *maximal* mismatch m , i.e.,

$$\mu^* \leq m , \quad (7.11)$$

everywhere in the template bank.

In the presence of noise, μ as defined in Eq. (7.8) is not directly accessible, and we therefore introduce a related quantity, namely the fractional loss of *measured* SNR^2 , namely

$$\check{\mu} \equiv \frac{2\mathcal{F}(\lambda_s) - 2\mathcal{F}(\lambda)}{2\mathcal{F}(\lambda_s) - 4} . \quad (7.12)$$

Note that $\check{\mu} \leq 1$, but contrary to Eq. (7.8) it can also be (slightly) negative, as we can have $2\mathcal{F}(\lambda_s) < 2\mathcal{F}(\lambda)$ due to noise.

For semicoherent searches, the metric is found [64] as the average of the fully coherent metrics over all the segments, namely

$$\widehat{g}_{ij}(\lambda) = \frac{1}{N} \sum_{k=1}^N g_{ij,k}(\lambda) , \quad (7.13)$$

where $\widetilde{g}_{ij,k}$ is the coherent metric, Eq. ((7.9)), in segment k .

A standard tool for parameter estimation is provided by the Fisher information matrix, which characterizes the statistical uncertainty of the maximum likelihood estimators (MLE) λ_{MLE}^i for the signal parameters λ_s^i . This can be formulated [104, 41, 7] as the well-known Cramer-R ao lower bound on the variance of an unbiased MLE (i.e. $E[\lambda_{\text{MLE}}^i] = \lambda_s^i$), namely

$$\sigma^2[\lambda_{\text{MLE}}^i] \geq \{\Gamma^{-1}\}^{ii} , \quad (7.14)$$

where the matrix $\{\Gamma^{-1}\}^{ij}$ denotes the inverse of the Fisher matrix Γ_{ij} , which is closely related (e.g., [41]) to the metric g_{ij} , namely

$$\Gamma_{ij} = \rho^2 g_{ij} . \quad (7.15)$$

A semicoherent search over N segments can be considered as N different measurements; thus the semicoherent Fisher matrix yields [?]

$$\widehat{\Gamma} = \sum_{k=1}^N \Gamma_{ij,k} . \quad (7.16)$$

Assuming constant SNR^2 for the different segments we can rewrite (7.16) in terms of the semicoherent metric, Eq. (7.13), namely

$$\widehat{\Gamma} = N \overline{\rho^2} \widehat{g}_{ij} \quad (7.17)$$

and thus

$$\{\widehat{\Gamma}^{-1}\}^{ij} = \frac{\widehat{g}^{ij}}{N \overline{\rho^2}} , \quad (7.18)$$

where $\widehat{g}^{ij}$ is the inverse matrix of $\widehat{g}_{ij}$.

7.1.3. Computing cost

The computing cost C of a fully coherent (or an ideal semicoherent [98]) search is primarily due to the computation of the $\mathcal{F}$ -statistic over all the templates. For a search over $\mathcal{N}$ templates using N segments of data from N_{det} detectors [98], the computing cost C is

$$C = N\mathcal{N}N_{\text{det}}c_1, \quad (7.19)$$

where c_1 is the implementation-dependent computing cost for a single template, segment and detector. A method of $\mathcal{F}$ -statistic computation based on short Fourier transforms (SFTs) [96] of length T_{SFT} is currently widely used in CW searches and will be considered in this chapter. The cost per template in this case is proportional to the segment duration, namely

$$c_1^{\text{SFT}} = c_0^{\text{SFT}} \frac{\Delta T}{T_{\text{SFT}}}, \quad (7.20)$$

where c_0^{SFT} is implementation- and hardware-dependent fundamental computing cost per SFT. Using the total number of SFTs

$$N_{\text{SFT}} = N\mathcal{N}N_{\text{det}} \frac{\Delta T}{T_{\text{SFT}}}, \quad (7.21)$$

we can write the total computing cost, Eq. (7.19), of the SFT-method as

$$C = \mathcal{N}N_{\text{SFT}}c_0^{\text{SFT}}. \quad (7.22)$$

In grid-based searches, the number of templates required to cover the search parameter space $\mathbb{P}$ is given by the general expression [54, 55]

$$\mathcal{N} \equiv \theta_n m^{-n/2} \int_{\mathbb{P}} d^n \lambda \sqrt{\det g}, \quad (7.23)$$

where θ is the normalized lattice thickness, n is the number of search dimensions, m is the maximal template-bank mismatch, Eq. (7.11), and $\det g$ is the determinant of the parameter-space metric, Eq. (7.9). The normalized thickness is a constant depending on the grid structure, e.g. for a hypercubic lattice $\theta_{\mathbb{Z}_n} = n^{n/2}2^{-n}$. The metric g_{ij} depends strongly on the duration ΔT and the number of segments N , in such a way that longer observation times typically require a (vastly) increased number of templates [53].

7.2. Coherent follow-up of semicoherent candidates

7.2.1. Basic two-stage search strategy

Here we introduce a simple two-stage strategy for following-up candidates from semicoherent searches. In the first stage, called *refinement*, we employ a finer search using the semicoherent statistic $\overline{2\mathcal{F}}$ to improve the initial maximum-likelihood estimator. In the

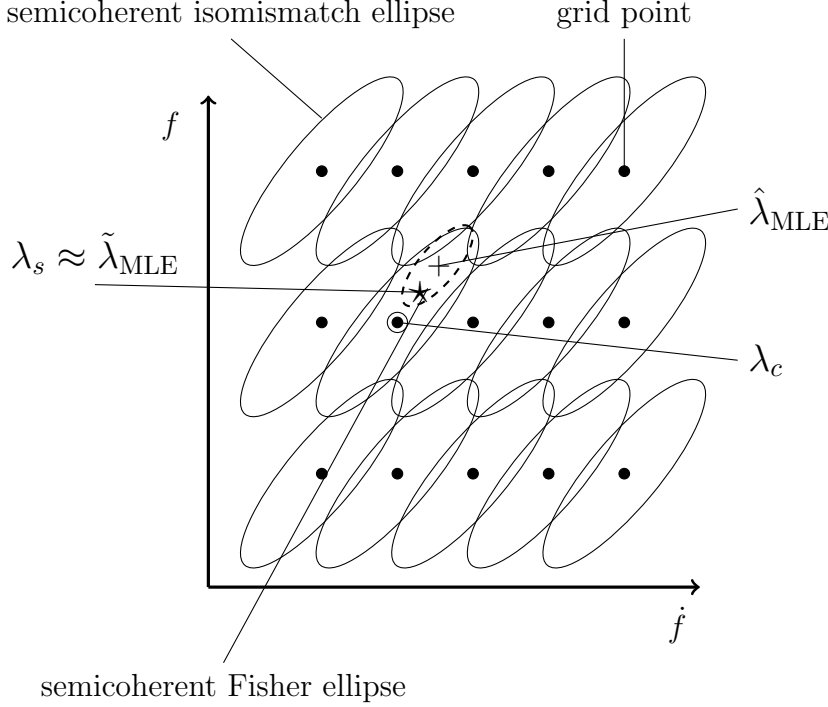


Figure 7.1.: Two-dimensional search grid in $\{f, \dot{f}\}$ space. The black dots are the search templates, placed such that the loss of SNR on any putative signal λ_s will be bounded by a maximal mismatch m , which defines the semicoherent isomismatch ellipses. The semicoherent Fisher ellipse centered on the MLE $\hat{\lambda}_{\text{MLE}}$ is used to constrain the zoom parameter space. The aim of the zoom stage is to find $\tilde{\lambda}_{\text{MLE}}$. Note that from Eqs. (7.17), (7.24) and (7.25) we see that the size of the metric ellipses and the Fisher ellipse is equal for $m = n_{\text{B}}^2 / N \bar{\rho}^2$.

second stage, called *zoom*, we apply the fully coherent statistic $\widetilde{2\mathcal{F}}$ using all the data T , in order to test whether the candidate is inconsistent with Gaussian noise and if it further agrees with the signal model.

The motivation for this two-stage approach can be seen from an example 2D search grid shown in Fig. 7.1. The search templates are generally placed such that a putative signal λ_s will be recovered with a loss of SNR bounded by a maximal mismatch m , as given in Eq. (7.11), namely

$$g_{ij} \Delta \lambda^i \Delta \lambda^j \leq m, \quad (7.24)$$

where equality defines an (n -dimensional) isomismatch ellipse. The initial semicoherent search will yield “candidates” $\hat{\lambda}_c$ for which the statistic $2\overline{\mathcal{F}}$ exceeds a certain threshold and is higher than neighboring templates.

The initial *refinement* stage of our follow-up strategy therefore consists in finding the (nearby) parameter-space point $\hat{\lambda}_{\text{MLE}}$ of the actual (local) maximum in the statistic

$2\overline{\mathcal{F}}(\hat{\lambda})$ (which is a smooth function of $\hat{\lambda}$), referred to as the maximum-likelihood estimator (MLE). This can be achieved simply by a denser placement of templates using the original statistic, i.e. by keeping the search setup unchanged in terms of the number and length of segments.

In the *zoom* stage, we fully coherently search the Fisher ellipse centered on the semi-coherent MLE $\hat{\lambda}_{\text{MLE}}$. This defines the parameter-space region that should contain the signal location λ_s with confidence corresponding to n_B standard deviations, i.e.,

$$\hat{\Gamma}_{ij} \delta\lambda^i \delta\lambda^j \leq n_B^2, \quad (7.25)$$

where $\delta\lambda^i = \hat{\lambda}_{\text{MLE}}^i - \lambda_s^i$. Note that the Fisher ellipse actually describes the fluctuations of the maximum likelihood estimator $\hat{\lambda}_{\text{MLE}}$ for a given signal location. However, provided the likelihood-manifold is not strongly curved, this also describes our uncertainty of the signal location for given MLE $\hat{\lambda}_{\text{MLE}}$, as indicated in Fig. 7.2. The zoom stage will yield the fully coherent maximum-likelihood estimator $\tilde{\lambda}_{\text{MLE}}$, which represents our best estimate for the signal parameters λ_s . Thus, the two-stage search strategy corresponds to the transition

$$\hat{\lambda}_c \xrightarrow{\text{refinement}} \hat{\lambda}_{\text{MLE}} \xrightarrow{\text{zoom}} \tilde{\lambda}_{\text{MLE}} \approx \lambda_s$$

In the following we use a subscript R to denote quantities in the refinement stage and a subscript Z for quantities in the zoom stage.

The search volume for the refinement stage depends on the template bank construction of the original semicoherent search. Ideally one iso-mismatch ellipse corresponding to the original template-bank construction (see Fig. 7.1) should be sufficient. In practice, however, it might often be necessary to use several grid spacings in each direction, if the template bank was not originally constructed in a strictly metric way or the metric approximation is not valid. In this case the exact number of grid spacings will have to be empirically determined in a Monte Carlo study.

Bounding box and volume of n -dimensional ellipses

In the following discussion, it will be useful to express the bounding box and volume of an n -dimensional ellipse, namely for the isomismatch ellipse of Eq. (7.24) and the Fisher ellipse of Eq. (7.25). The general form of the n -dimensional ellipse equation is

$$G_{ij} d\lambda^i d\lambda^j = R^2, \quad (7.26)$$

where G_{ij} is a positive-definite symmetric matrix. The extents $\Delta\lambda^i$ of a bounding box along coordinate axes λ^i (as indicated in Fig. 7.2) can be obtained from the diagonal elements of the inverse matrix, $\{G^{-1}\}^{ij}$, namely

$$\Delta\lambda^i = 2R \sqrt{\{G^{-1}\}^{ii}}. \quad (7.27)$$

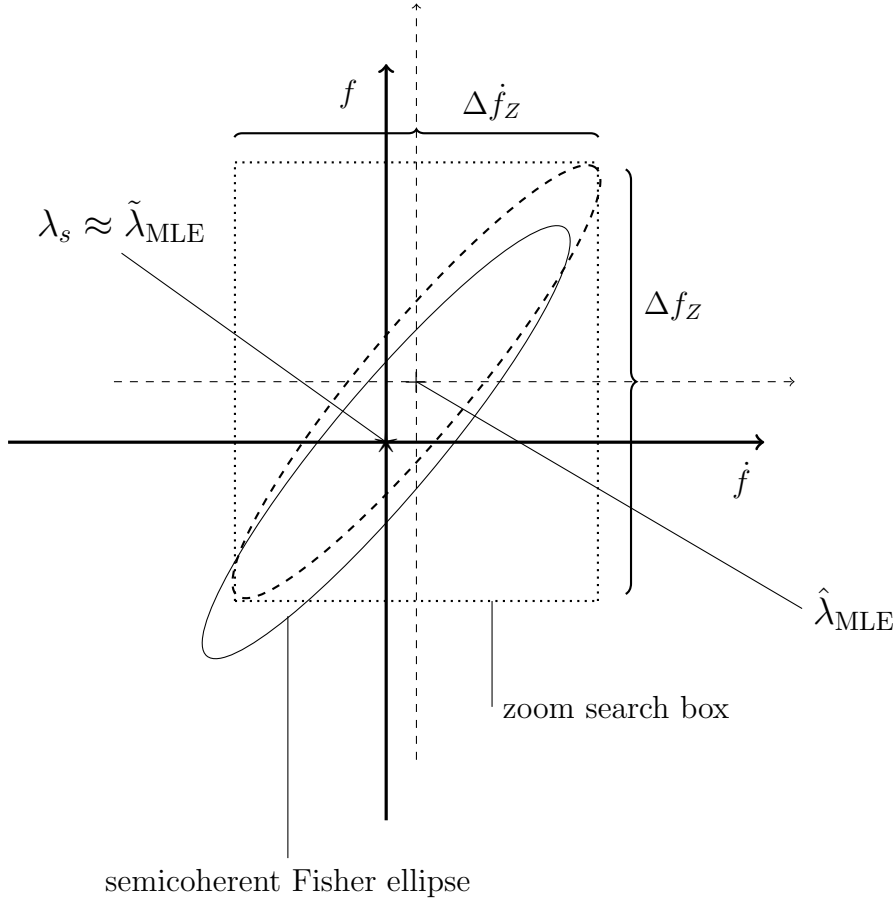


Figure 7.2.: Two-dimensional example: Fisher ellipse, Eq. (7.25), defining the zoom search space, centered on the semicoherent MLE $\hat{\lambda}_{\text{MLE}}$. The extents $\{\Delta f, \Delta \dot{f}\}$ of the bounding box are given by Eq. (7.27).

The ellipse coordinate volume is expressible via the matrix determinant, $\det G$, namely

$$V = \frac{R^n}{\sqrt{\det G}} \mathcal{V}_n, \quad (7.28)$$

where $\mathcal{V}_n = \frac{\pi^{n/2}}{\Gamma(1+n/2)}$ is the volume of unit n -ball.

7.2.2. Classification of zoom outcomes

Assuming a real CW signal, we can estimate the range of expected values of the fully coherent zoom $\mathcal{F}$ -statistic in $\tilde{\lambda}_s$. From Eq. (7.5) we can obtain a (rough) estimate of the average SNR^2 from the *measured* average SNR^2 of the semicoherent maximum likelihood estimator, namely

$$\overline{\rho^2}_{\text{MLE}} \approx \overline{2\mathcal{F}}_{\text{MLE}} - 4. \quad (7.29)$$

The SNR^2 of the fully coherent search is linear in the number of segments N , i.e.,

$$\tilde{\rho}^2 = N \overline{\rho^2}_{\text{MLE}}. \quad (7.30)$$

Substitution of the above expression into Eq. (7.2) yields the expectation for the fully coherent matched filter in $\tilde{\lambda}_{\text{MLE}}$, namely

$$\widetilde{2\mathcal{F}}_o \equiv E[\widetilde{2\mathcal{F}}] \approx 4 + N \overline{\rho^2}_{\text{MLE}}. \quad (7.31)$$

Further substitution of Eq. (7.30) into Eq. (7.3) yields the corresponding variance as

$$\sigma_o^2 \equiv \sigma^2[\widetilde{2\mathcal{F}}] \approx 2(4 + 2N \overline{\rho^2}_{\text{MLE}}). \quad (7.32)$$

These quantities are useful for defining what we mean by confirming a CW signal.

Note that the uncertainty in the original SNR-estimation in Eq. (7.29) results in a distribution around the final estimate of Eq. (7.31) that is wider than estimated by Eq. (7.32). This effect can be computed analytically and empirically, and is found to amount to about a factor of 2.

Depending on the maximal $\widetilde{2\mathcal{F}}$ value found in the final zoom stage, we can distinguish three possible outcomes:

- *Consistency with Gaussian noise (G).* - The fully coherent $\widetilde{2\mathcal{F}}$ value does not exceed a threshold

$$\widetilde{2\mathcal{F}} < G_{\text{th}}, \quad (7.33)$$

where G_{th} is chosen to correspond to some (small) false-alarm probability p_{fA} in Gaussian noise.

For example, a threshold $G_{\text{th}} = 60$ corresponds to a very small false-alarm probability of order 10^{-12} in a single template, as given by Eq. (7.43).

- *Non-Gaussian origin ($\neg G$).* - The candidate is loud enough to be inconsistent with Gaussian noise at the false-alarm probability p_{fA} , i.e.

$$\widetilde{2\mathcal{F}} \geq G_{\text{th}} . \quad (7.34)$$

- *Signal recovery (S).* - A *subclass* of $\neg G$; the final zoomed candidate $\widetilde{2\mathcal{F}}$ exceeds the Gaussian noise threshold G_{th} and falls into the predicted signal interval given by Eqs. (7.31) and (7.32) (at some confidence level). We can write this as

$$\widetilde{2\mathcal{F}}_{\text{th}}^{(S)} < \widetilde{2\mathcal{F}} < \widetilde{2\mathcal{F}}_{\text{max}}^{(S)} \quad (7.35)$$

where $\widetilde{2\mathcal{F}}_{\text{th}}^{(S)} \equiv \max\{G_{\text{th}}, \widetilde{2\mathcal{F}}_o - n_u \sigma_o\}$, and $\widetilde{2\mathcal{F}}_{\text{max}}^{(S)} \equiv \widetilde{2\mathcal{F}}_o + n_u \sigma_o$, where n_u determines the desired confidence level. Here we consider $n_u = 6$, which corresponds roughly to a confidence of $\sim 99.6\%$.

Note that there can be cases where a zoomed candidate ends up in $\neg G$ but does not make it into the signal recovery (S) band, e.g. typically $G_{\text{th}} < \widetilde{2\mathcal{F}} < \widetilde{2\mathcal{F}}_{\text{th}}^{(S)}$. There can be different reasons for this, e.g. the search algorithm converged to a secondary maximum in the refinement or zoom stage, the signal model deviates from reality and requires modification, or the “signal” found is of nonastrophysical origin (e.g., a detector-noise artifact). Generally further investigation will be required for all candidates falling into the non-Gaussian category ($\neg G$).

7.2.3. Grid-based computing cost of the zoom stage

We do not consider a grid-based follow-up method in this chapter, but it is instructive to estimate the corresponding computing cost for later comparison. To estimate the number of templates required for the fully coherent search, we can use Eq. (7.28) to compute the volume of the follow-up Fisher ellipse, Eq. (7.25), and divide it by the volume covered by one coherent template, Eq. (7.24). Namely, the Fisher ellipse volume is given by

$$\widehat{V} = \frac{n_{\text{B}}^n}{(N \overline{\rho^2})^{n/2} \sqrt{\det \widehat{g}}} \mathcal{V}_n , \quad (7.36)$$

while the coherent template volume at mismatch m is

$$\widetilde{V} = \frac{m^{n/2}}{\sqrt{\det \widetilde{g}}} \mathcal{V}_n . \quad (7.37)$$

Therefore, we can estimate the number of templates as

$$\mathcal{N} \approx \frac{\widehat{V}}{\widetilde{V}} = \frac{n_{\text{B}}^n}{(N \overline{\rho^2})^{n/2} m^{n/2}} \frac{\sqrt{\det \widetilde{g}}}{\sqrt{\det \widehat{g}}} . \quad (7.38)$$

Consider a follow-up of a candidate from a directed $n = 2$ search in $\{f, \dot{f}\}$ (e.g. see Fig. 7.1). Assuming a semicoherent search using $N = 200$ segments of $\Delta T = 1$ days duration without gaps, and a fully coherent observation time of $T = 200$ days. Using the expressions found in [56], the determinants of the two-dimensional coherent and the semicoherent metrics are found as

$$\sqrt{\det \tilde{g}} = \pi^2 T^3 \frac{1}{540}, \quad (7.39)$$

$$\sqrt{\det \hat{g}} = \pi^2 \Delta T^3 \frac{\gamma(N)}{540}, \quad (7.40)$$

where $\gamma \approx \sqrt{5} N$ is the spindown refinement factor. Putting everything together in Eq. (7.38), we obtain

$$\mathcal{N} \approx \frac{n_B^2 N}{\sqrt{5} \bar{\rho}^2 m}, \quad (7.41)$$

where we used $N = T/\Delta T$. For a signal with $\bar{\rho}^2 = 1$, $n_B = 24$ ¹ and $m = 0.1$ the number of templates is therefore $\mathcal{N} \approx 5.1 \times 10^5$. Thus, using Eq. (7.22) for two detectors and the SFT method with $T_{\text{SFT}} = 1800$ s, the zoom computing cost is $C \approx 11$ min per candidate, where we used the fundamental computing cost $c_0^{\text{SFT}} = 7 \times 10^{-8}$ s[98].

In the more general case where the sky position of the source is also unknown, the number of sky points typically scales at least quadratically with the observation time [112, 56] (for coherent integration longer than few days), thus generally resulting in completely prohibitive computational requirements for grid-based follow-up searches. In particular, extending the directed search example from the previous paragraph to an all-sky follow-up would require $\mathcal{N}_{\text{sky}} \approx 1.3 \times 10^6$ sky points², or a total of $\mathcal{N} \approx 6.8 \times 10^{11}$ templates.

For comparison, using the gridless search algorithm discussed in the next sections, it is possible to coherently follow up 2D directed candidates in less than 2 minutes (see Fig. 7.4(d)), and all-sky candidates in about 1 hour per candidate (see Fig. 7.5(d)).

7.2.4. Mesh Adaptive Direct Search (MADS)

A significant difference between the hierarchical search strategies discussed in [53, 64, 88] and in this approach is the method of template bank construction at the different stages. Namely, we consider a grid-less method for exploring the parameter space.

The MADS class of algorithms for derivative-free optimization has been first introduced in [70] and further developed in [71, 72] among others. In this subsection we only introduce some of the control parameters of the algorithm required in the construction

¹This large n_B value is found to contain the signal location in more than 98 % of the cases even for weak signals, where the Fisher-matrix may be a poor predictor; see [104].

²The number of sky templates has been estimated by numerical computation of the sky part of the metrics $\hat{g}$ and $\tilde{g}$ using `FstatMetric.v2` from LALSUITE [105], see also [41].

of MADS-based $\mathcal{F}$ -statistic searches; for an in-depth treatment and proofs we refer the reader to the cited publications.

MADS consist of the iteration of two steps, called *search* and *poll*, in which trial points are constructed and evaluated in order to find an extremum. In the search step, any strategy can be applied to construct trial points. We use quadratic models (quadratic form) to approximate the objective function from a sample of objective values [72]. If the local exploration in the search step fails to generate a new solution, a set of poll points is generated using a stochastic or deterministic method. Stochastic means that the poll points are generated randomly [70], where as deterministic refers to the usage of pseudo random Halton sequences [71]. However, both methods generate points in directions, which form a dense set in the unit sphere after an infinite number of iterations. For a given starting point λ_c with parameter-space boundaries $\Delta\lambda_B$, initial step sizes $d\lambda$ and a method for generation of poll points, the discretization of the parameter space Δ_k^m at iteration k is governed by a fixed rational number $u_b > 1$ and the coarsening $w^+ \geq 0$ and refining $w^- \leq -1$ exponents. If the current iteration generates a better solution, the discretization in the next iteration is coarser, namely $\Delta_{k+1}^m = u_b^{w^+} \Delta_k^m$, otherwise $\Delta_{k+1}^m = u_b^{w^-} \Delta_k^m$ [70]. The algorithm stops if an improved solution cannot be found or the total number of evaluated parameter space points p reaches some given maximum $p_{\max}$.

7.2.5. MADS-based follow-up algorithm

From the point of view of the MADS algorithm, the function to optimize is a blackbox requiring some input to produce a single output value. The blackbox in our case is either the computation of the semicoherent $\mathcal{F}$ -statistic $2\overline{\mathcal{F}}$ of Eq. (7.4) in the refinement, or the fully coherent $\mathcal{F}$ -statistic $\widetilde{2\mathcal{F}}$ in the zoom stage. In order to minimize the possibility of convergence to secondary maxima, we run multiple instances of the MADS search in each stage, varying the mesh-coarsening exponent w^+ . The minimal $w_{\min}^+$ and maximal $w_{\max}^+$ coarsening exponents determine the number of MADS steps in each pass, namely $n_{\text{steps}} = w_{\max}^+ - w_{\min}^+ + 1$. Thus, we consider our search algorithm to be composed of several instances of MADS; see Fig. 7.3. The inputs of the search algorithm are the candidate λ_c to follow-up, the search boundaries $\Delta\lambda_{R/Z}$ around the candidate, and a set of MADS input parameters, namely $\{d\lambda, u_b, w_{\min}^+, w_{\max}^+, w^-\}$. In the zoom stage, the search boundaries ($\Delta\lambda_Z$) are estimated from the bounding box of the Fisher ellipse, using Eq. (7.27). For the refinement stage, the search boundaries ($\Delta\lambda_R$) generally have to be determined depending on the templatebank setup of the original semicoherent search. Note, however, that the bounding boxes $\Delta\lambda$ only serve as a necessary input parameter to the NOMAD search algorithm, while the effective MADS search region can be further reduced by rejecting points that do not satisfy a given constraint. For example, the effective search region in the zoom stage always consists of the Fisher ellipse, Eq. (7.25).

The initial step sizes $d\lambda^i$ are also empirically determined, typically as some fraction of the search boundary $\Delta\lambda_{R/Z}^i$.

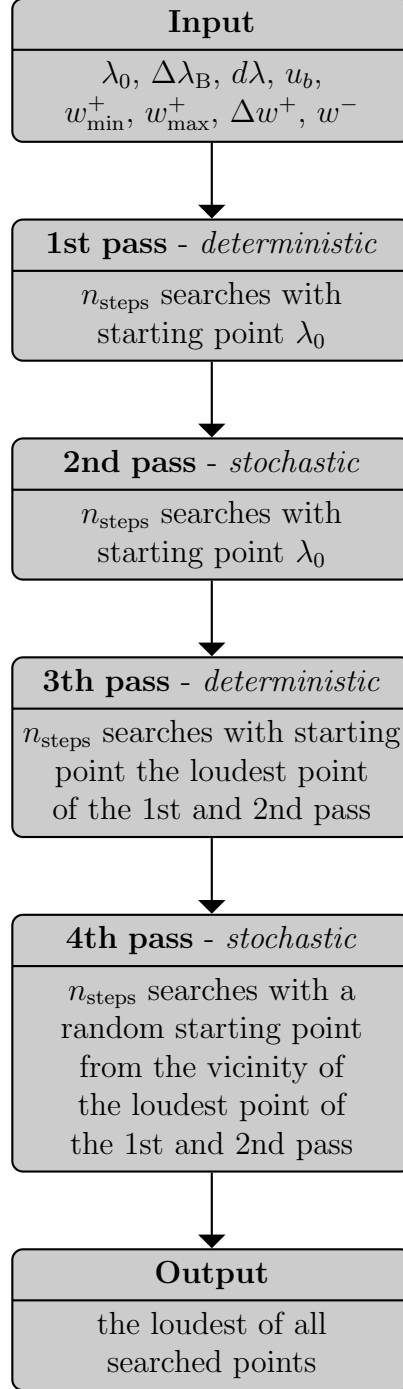


Figure 7.3.: MADS-based search algorithm with four passes, where $\lambda_0 = \lambda_c$ in the refinement stage and $\lambda_0 = \hat{\lambda}_{\text{MLE}}$ in the zoom stage.

We propose a four pass algorithm with an equal (for simplicity) number of steps n_{steps} in each pass; however, with different starting point and method of trial-point generation:

- *1st pass.* - Starting point λ_c , *deterministic* point generation,
- *2nd pass.* - Starting point λ_c , *stochastic* point generation,
- *3rd pass.* - Starting point loudest template from the first two passes, *deterministic* point generation,
- *4th pass.* - Starting point from the vicinity of the loudest point from the first two passes, *stochastic* point generation.

In the zoom stage, we terminate the search as soon as the loudest point of the current iteration satisfies the signal-confirmation condition (S) of Eq. (7.35). In lower-dimensional cases, such as the directed search considered later, a single pass is therefore often found to be sufficient. For later usage, we introduce the total number of MADS iterations n_I as the sum of the number of steps in each pass.

7.2.6. MADS follow-up computing cost

Contrary to grid-based searches, the computing cost of the MADS based algorithm is nondeterministic, due to the *a priori* unknown number of explored parameter-space points. To estimate the maximal computing cost of the refinement or the zoom stage using Eq. (7.19), we need the maximal number of possibly evaluated templates

$$\mathcal{N}_{\text{max}} = \sum_{i=0}^{n_I} p_{\text{max}}^i, \quad (7.42)$$

where p_{max}^i is the user-specified maximum of the number of computed templates at MADS iteration i . This maximal number is typically chosen to be large to avoid too-early interruption of the MADS instance, e.g., when further improvement of the current solution is possible while the extremum is not yet found. However, if the extremum is found, a MADS iteration starting from this point terminates rapidly.

Note that the fundamental computing cost c_0^{SFT} in stochastic searches over the sky is typically larger than in a grid-based search, where a lot of templates with different spindown components can be computed at fixed sky position. This results in a larger value of about $c_0^{\text{SFT}} \approx 3 \times 10^{-7}$ s instead of the number quoted in Sec. 7.2.3.

7.2.7. False-alarm and detection probability

After the final fully coherent zoom stage we are left with a candidate falling into one of the three categories discussed earlier: namely, the candidate is consistent with the signal model (S), with Gaussian noise (G), or is of non-Gaussian origin but inconsistent

with the signal model. An additional valuable piece of information is the false-alarm probability associated with the candidate. This is the probability of exceeding a threshold $2\mathcal{F}$ value in the absence of a signal, where the relevant distribution is the central χ^2 distribution with four degrees of freedom, denoted as $\chi_4^2(2\mathcal{F})$. The single-template false-alarm probability is

$$\begin{aligned} p_{\text{fA}}^1 &= \int_{2\mathcal{F}_{\text{th}}}^{\infty} d(2\mathcal{F}) \chi_4^2(2\mathcal{F}) \\ &= (1 + \mathcal{F}_{\text{th}}) e^{-\mathcal{F}_{\text{th}}} , \end{aligned} \quad (7.43)$$

and for $\mathcal{N}$ independent templates, this results in

$$p_{\text{fA}} = 1 - (1 - p_{\text{fA}}^1)^{\mathcal{N}} , \quad (7.44)$$

where for $\mathcal{N} p_{\text{fA}}^1 \ll 1$, Taylor expansion yields $p_{\text{fA}} \approx \mathcal{N} p_{\text{fA}}^1$. For example, a threshold of $G_{\text{th}} = 70$ for a search with $\mathcal{N} = 1 \times 10^5$ templates corresponds to a false-alarm probability of $p_{\text{fA}} \lesssim 2 \times 10^{-9}$, where the upper bound corresponds to $\mathcal{N}$ completely independent templates.

The overall detection probability of the follow-up method depends on the signal SNR. Higher SNR in the refinement stage yields better localization of the signal - i.e., a smaller Fisher ellipse - and thus also a higher probability of signal recovery, Eq. (7.35). In addition, the MADS-algorithm parameters also affect the detection efficiency, e.g. an increased number of MADS iterations increases the detection probability, especially for signals with lower SNR. Because of this, the detection probability will have to be estimated empirically in a Monte Carlo study (see Figs. 7.4(c) and 7.5(c)).

7.3. Monte Carlo studies

To demonstrate the capability of the systematic follow-up procedure proposed in Sec. 7.2, we perform two different types of Monte Carlo (MC) studies.

In the first case, we simulate a so-called *directed* search for a fixed sky position, where we follow up candidates in a two-dimensional spindown space, i.e. $\{f, \dot{f}\}$. In the second case, we simulate an all-sky search over the four-dimensional parameter space $\{\alpha, \delta, f, \dot{f}\}$.

All MADS searches are implemented using the MADS reference library NOMAD [101], and the LAL library from the LALSUITE [105] is used for the $\mathcal{F}$ -statistic computation [44]. The Gaussian data and signal injections are produced using the LALAPPS programs from LALSUITE. In particular, with `lalapps_Makefakedata_v4`, we create data sets of total duration $T = 200$ days, with $N = 200$ segments of duration $\Delta T = 1$ days, using SFTs of length $T_{\text{SFT}} = 1800$ s, for the two LIGO detectors H1 and L1. The noise level per detector is generated as Gaussian white noise with a power-spectral density S of $\sqrt{S} = 2 \times 10^{-23} \text{ Hz}^{-1/2}$.

Independently of the type of search, the initial candidates to follow-up are prepared as follows. Rather than performing a semicoherent grid-based search using the Hough [43]

or GCT method [92], we generate candidates by drawing a random point in the vicinity of the injection and consider it a candidate if the semicoherent metric mismatch μ^* is within the range

$$\mu^* \in [0, 1] ; \quad (7.45)$$

see Figs. 7.4(a) and 7.5(a). This procedure for candidate preparation allows us to separate the study of the follow-up algorithm from the problem of how to setup a semicoherent search, which is a difficult question on its own.

Note that even if the original grid-based semicoherent search does not produce candidates that conform with Eq. (7.45), we can always increase the density of the grid until Eq. (7.45) applies. This would amount to a (cheap) pre-processing stage inserted before the present follow-up procedure .

7.3.1. Follow-up of candidates from a directed search

For the directed type of searches, we fix the sky position to the coordinates of the Galactic Center. This choice is arbitrary and we could use any other point without qualitatively changing the results. We create 5000 data sets. Note that each data set has different Gaussian noise realization in which a CW signal from an isolated source is injected. In the process of injection, the original noise data set is also used to examine the behavior of the follow-up method in the absence of a signal.

The pulsar injection parameters λ_s are drawn uniformly in the range $f \in (50, 51)$ Hz, $\cos \iota \in (-1, 1)$, $\psi \in (-\pi/4, \pi/4)$, and $\phi_0 \in (0, 2\pi)$, where ι is the inclination angle of the source with respect to the line-of-sight, ψ is the polarization, and ϕ_0 the initial phase of the signal [44]. The signal amplitude h_0 is chosen such that the expected average SNR² of Eq. (7.7) for a perfect match is distributed uniformly in the range $\bar{\rho}_s^2 \in (0, 2)$. The spindown $\dot{f}$ is chosen uniformly in the range $\dot{f} \in (-\frac{f_{\min}}{\tau_{\min}}, \frac{f_{\min}}{\tau_{\min}})$ with a minimal spindown age $\tau_{\min} = 300$ y at $f_{\min} = 50$ Hz. The MADS algorithm parameters used in the MC , which have been found empirically to achieve good results are summarized in Table 8.1. For this type of follow-up, we find that the first pass of the search algorithm in

stage	w^-	$w_{\min}^+$	$w_{\max}^+$	u_b	$p_{\max}$
R	-1	1	1	2	20000
Z	-1	1	50	1.1	20000

Table 7.1.: Algorithm parameters for follow-up of candidates from directed searches.

the refinement stage and only two repetitions of the second pass in the zoom stage are sufficient. We restrict the size of the search box for the refinement stage $\Delta\lambda_R$ by taking one frequency and two first spindown metric extents. In the zoom stage we constrain the parameter space to a Fisher ellipse Eq. (7.25) with $n_B = 24$.

We first apply the follow-up chain to the pure Gaussian noise data without injected signals. The corresponding $2\mathcal{F}_Z$ distribution of the resulting fully coherent zoom stage is plotted in Fig. 7.4(b). The maximal value found is $2\mathcal{F}_Z^{\max} = 51.61$. We therefore use

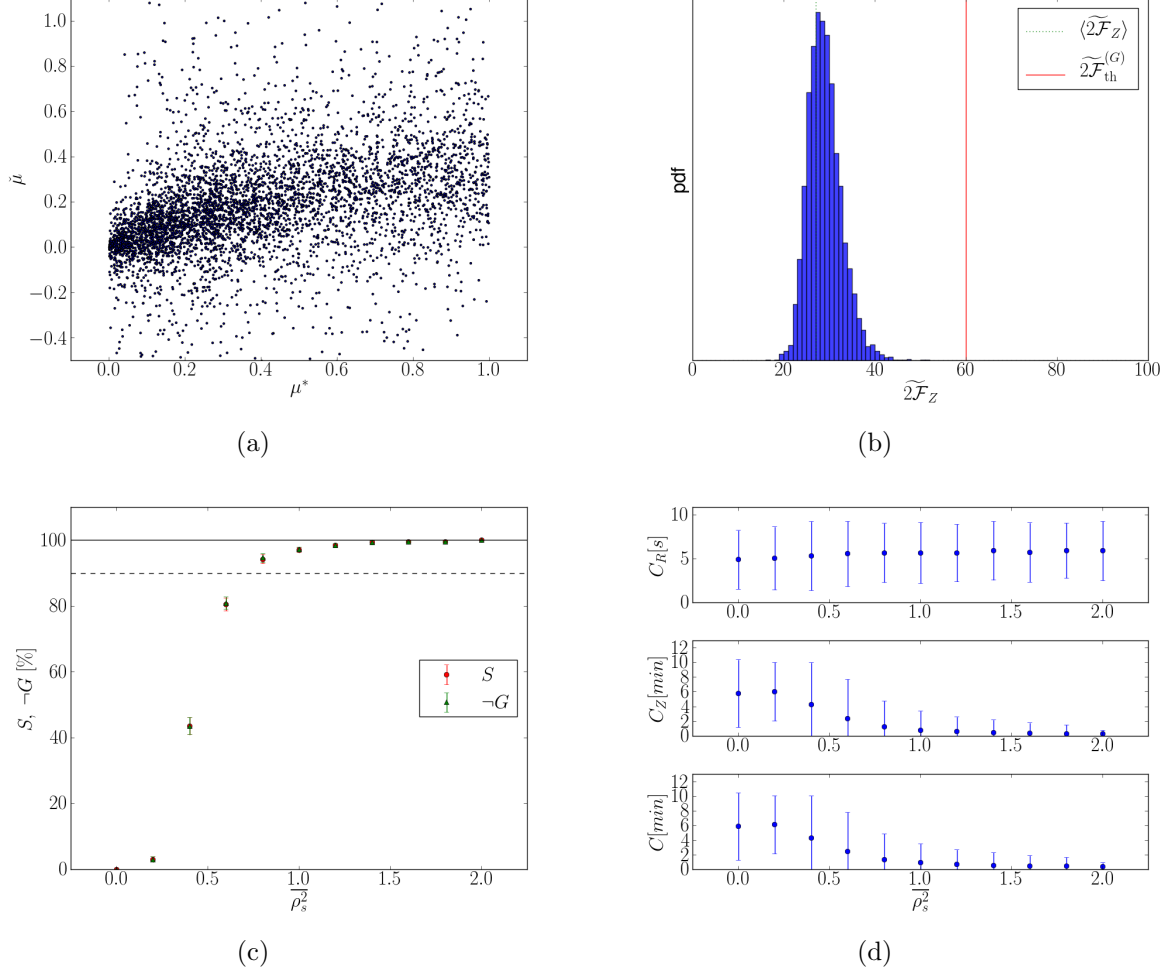


Figure 7.4.: Monte Carlo study of 2-stage follow-up of candidates from a directed $\{f, \dot{f}\}$ semicoherent search pointed toward the Galactic Center with $N = 200$ segments of duration $\Delta T = 1$ days. (a) SNR loss of the initial candidates $\tilde{\mu}$ versus semicoherent metric mismatch μ^* to the closest template. (b) $\tilde{\mathcal{F}}_Z$ distribution after the fully coherent 2-D $\{f, \dot{f}\}$ zoom stage of 5000 directed searches in pure Gaussian noise without injected signal. The maximal $2\mathcal{F}$ -value found is $\tilde{\mathcal{F}}_Z^{\text{max}} = 51.61$. The mean value $\langle \tilde{\mathcal{F}}_Z \rangle = 29.00$ is plotted with a dotted line. (c) Percentage of the 5000 injected signals classified as recovered (S) and of non-Gaussian origin ($\neg G$) as function of the non-centrality parameter $\bar{\rho}_s^2$, Eq. (7.7). The error bars are computed by using a Jackknife [113] estimator. (d) *Upper plot*: computing cost of the semicoherent refinement stage. *Middle plot*: computing cost of the fully coherent zoom stage. *Lower plot*: total computing cost.

a threshold for the classification of non-Gaussian candidates ($\neg G$) of $G_{\text{th}} = 60$, which is safely above this level.

We next apply the follow-up chain to the Gaussian noise data with injected signal.

In Fig. 7.4(c) we plot the percentage of injected signals that are classified as recovered signals (S) and non-Gaussian origin ($\neg G$) as a function of the injected signal strength $\bar{\rho}_s^2$. From this plot, we can read out the detection probability - namely, we reach 90% of signal recovery for candidates with $\bar{\rho}_s^2 \approx 0.7$.

The computing cost as a function of $\bar{\rho}_s^2$ is plotted in Fig. 7.4(d). We notice that the cost of the refinement stage is negligible, and in the zoom stage, the averaged computing time decreases with higher signal strength.

7.3.2. Follow-up of candidates from an all-sky search

The data and signal preparation for the following all-sky Monte Carlo study is the same as in the directed search case; however, the sky position is drawn isotropically over the whole sky. We create 7500 data sets with uniformly distributed injected average SNR² in the range $\bar{\rho}_s^2 \in (0, 3)$. The algorithm parameters used in the refinement and zoom stage, which have been found empirically to yield good performance are given in Table 7.2. We also find that here the zoom stage benefits from performing all four search passes shown in Fig. 7.3. The size of the search box for the refinement stage in the spindown subspace has been defined exactly as in the directed search example. The sky subspace is constrained by using an $m = 1$ isomismatch ellipse. As in the previous example, we use $n_B = 24$ in Eq. (7.25) to determine the size of the Fisher ellipse.

stage	w^-	$w_{\min}^+$	$w_{\max}^+$	u_b	p
R	-1	1	5	2	20000
Z	-1	1	50	1.2	20000

Table 7.2.: Follow-up algorithm parameters for full parameter space searches.

Similarly to the directed follow-up, we first test the pipeline using the Gaussian noise data without injections. The resulting distribution of final $2\tilde{\mathcal{F}}_Z$ values is plotted in Fig. 7.5(b). The maximal value found is $\widetilde{2\mathcal{F}}_Z^{\max} = 58.76$, which is higher compared to the value found in the directed follow-up searches due to the increased number of evaluated templates. We therefore use a threshold for the classification of non-Gaussian candidates ($\neg G$) of $G_{\text{th}} = 70$, which is safely above this level.

Next, we search the data containing the injected signals. In Fig. 7.5(c) we plot the fraction of signals classified as recovered (S) and the percentage of MC trials found to be of non-Gaussian origin ($\neg G$) as a function of the injected signal strength $\bar{\rho}_s^2$. In order to achieve 90% signal recovery (S), we now need stronger signals, namely $\bar{\rho}_s^2 \gtrsim 1.7$. However, for $\bar{\rho}_s^2 \approx 1.5$, we can already achieve 90% “detection probability” in the sense of separating candidates from Gaussian noise ($\neg G$). This indicates that the zoom step sometimes converges on a secondary maximum. Given that any non-Gaussian

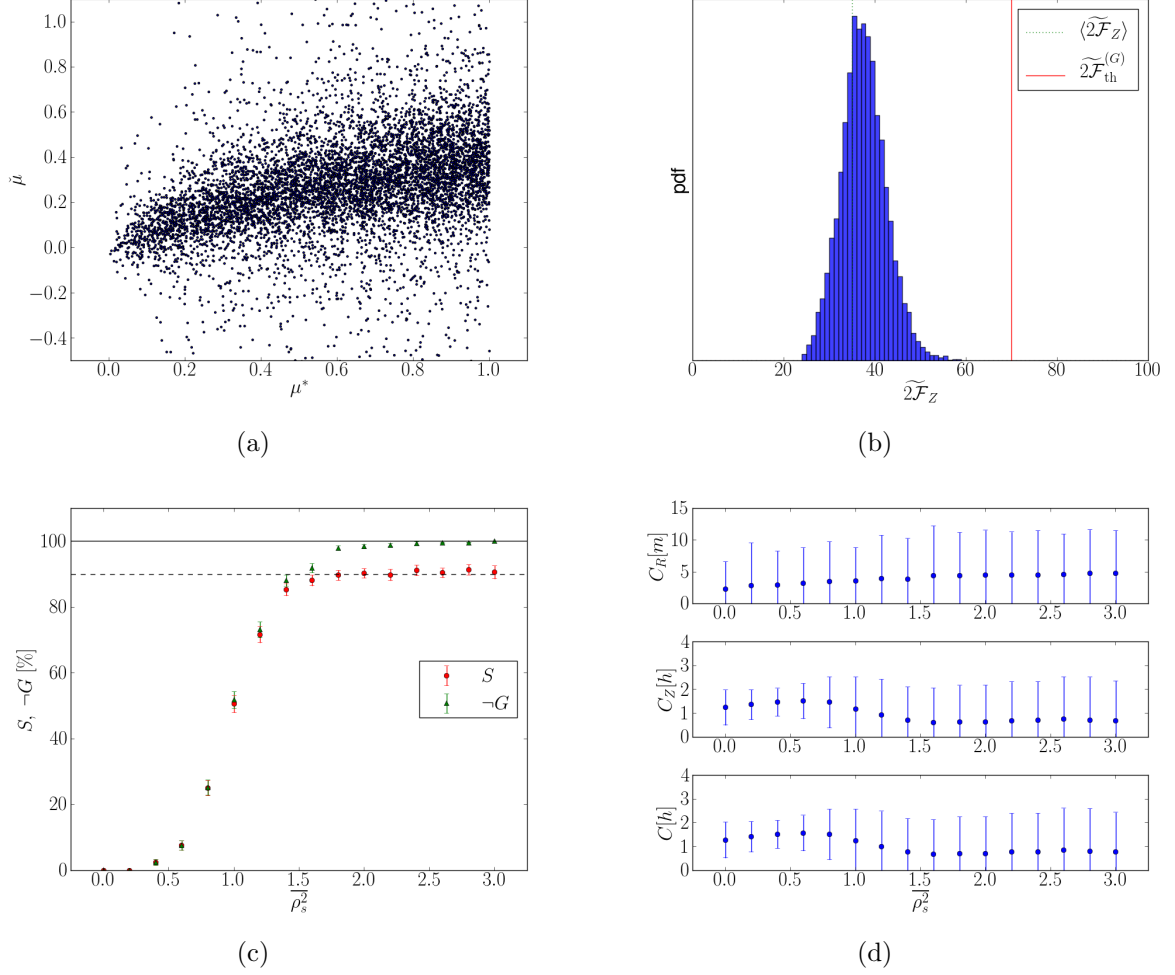


Figure 7.5.: Monte Carlo study of 2-stage follow-up of candidates from an all-sky $\{\alpha, \delta, f, \dot{f}\}$ semicoherent search with $N = 200$ segments of duration $\Delta T = 1$ days. (a) SNR loss of the initial candidates $\tilde{\mu}$ versus semicoherent metric mismatch μ^* . (b) $2\mathcal{F}_Z$ distribution after the fully coherent 4-D $\{\alpha, \delta, f, \dot{f}\}$ zoom stage of 7500 searches in pure Gaussian noise, without injected signal. The maximal $2\mathcal{F}$ -value found is $2\mathcal{F}_Z^{\text{max}} = 58.76$. The mean value is $\langle 2\mathcal{F}_Z \rangle = 37.50$ indicated with dots. (c) Percentage of the 7500 injected signals classified as recovered ($-S$) and of non-Gaussian origin ($\times \neg G$) as function of the signal strength ρ_s^2 . The error bars are computed using a Jackknife [113] estimator. (d) *Upper plot*: computing cost of the semicoherent refinement stage. *Middle plot*: computing cost of the fully coherent zoom stage. *Lower plot*: total computing cost.

($\neg G$) candidates after zoom will receive further scrutiny, it would be straightforward to further explore the parameter space around such candidates to localize a potential primary maximum.

The computing cost as a function of $\overline{\rho_s^2}$ is plotted in Fig. 7.5(d). We notice that the total computing cost is dominated by the zoom stage and the averaged computing time is rather independent of the signal strength.

7.4. Discussion

We have studied a two-stage scheme for the fully coherent follow-up of semicoherent candidates. The first stage, called refinement, aims to find the maximum-likelihood estimator of the initial semicoherent candidate. This allows us to better constrain the parameter space for the coherent zoom stage. The two-stage scheme is suitable for following-up candidates from all-sky or directed semicoherent searches. The proposed grid-less optimization lowers the computing cost per candidate to acceptable levels. In Monte Carlo studies we tested the efficiency of the algorithm for directed and all-sky follow-up searches.

In this chapter we restricted the all-sky follow-up optimization to four dimensions, namely sky, frequency and first spindown. Further work is required to extend the optimization to higher dimensions. A related attractive direction for further development is the extension and application of the search algorithm for follow-up of CW candidates in binary systems, which is a challenging higher-dimensional problem.

We also aim to extend the two-stage scheme presented here by including intermediate semicoherent zoom stages. This should allow us to further reduce the computing cost and increase detection efficiency.

8. Fully coherent follow-up of candidates from the Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data

As discussed in Chap. 2 the search for unknown sources of continuous gravitational waves (CWs) is computationally bound due to the enormous parameter space that needs to be covered [64]. Advanced semicoherent search techniques [88] [92] and a distributed computing environment such as Einstein@Home [102] are typically used to identify interesting regions of the parameter space, which then require fully coherent follow-up studies in order to confirm or discard potential gravitational-wave candidates. Although the parameter space associated with these candidates is greatly reduced, it is still large enough to lead to a prohibitive computing cost, when data of order of months or years is analyzed fully coherently with a classical grid-based method. However as we demonstrated in the previous chapter a follow-up method that combines the $\mathcal{F}$ -statistic [28, 46] with a Mesh Adaptive Direct Search (MADS) [70] based algorithm allows us to fully coherently examine long data sets at feasible computational cost.

In this chapter we adapt the two-stage algorithm proposed in Chap. 7 to follow-up some loud CW candidates identified in the recent Einstein@Home all-sky search for periodic gravitational waves in LIGO fifth Science Run (S5) data [59]. We first validate the follow-up pipeline in Monte Carlo studies by injecting and searching for simulated CW signals added into simulated Gaussian noise data. Then we apply the search method to 27 loud candidates identified with the Einstein@Home search and associated with 3 simulated signals (hardware injections). We also investigate 9 loud candidates identified with the Einstein@Home search and for which there is no known instrumental evidence.

This chapter is organized as follows: in Sec. 8.1 we summarize the original Einstein@Home LIGO search in S5 data, and in Sec. 8.2 we summarize the two-stage follow-up method and introduce the search pipeline. The efficiency of the follow-up algorithm is tested in Monte Carlo studies presented in Sec. 8.3. In Sec. 8.4 we present the follow-up results for the 27 candidates representing the 3 hardware injections. In Sec. 8.5 we show the results of the follow-up for the 9 loud CW outliers. We discuss in Sec. 8.6.

Notation

In the following we use the notation of [59] and the previous chapters, i.e., we use tilde for quantities Q when referring to a fully coherent stage, $\tilde{Q}$, and overhat when referring to a semicoherent stage, $\hat{Q}$. A quantity measured after the refinement stage is denoted with a subscript R , Q_R , and after the zoom stage with subscript Z , Q_Z .

8.1. The LIGO S5 all-sky search with Einstein@Home

The Einstein@Home all-sky search for periodic gravitational waves as described in [59] investigates a frequency range $f \in [50, 1190]$ Hz and a spindown range $\dot{f} \in [-20, 1.1] \times 10^{-10}$ Hz s⁻¹. The search uses the semicoherent Hough-transform method [43], where the entire amount of data is divided into N shorter segments of duration ΔT . In a first step a fully coherent $\mathcal{F}$ -statistic search is performed on a coarse grid for each of the segments. Then the Hough-statistic is computed on a finer grid from the $\mathcal{F}$ -statistic values of the individual segments. The Einstein@Home search has been split into 2 parts: the analysis of the first (second) year of S5, referred to as S5R3 (S5R5). In this chapter we consider only the S5R5 analysis, which spans data collected between the GPS times 852443819 s (Wed Jan 10 05:56:45 GMT 2007) and 875278812 s (Mon Oct 01 12:59:58 GMT 2007), or approx. 264 days [59]. The data collected by the Hanford (H1) and Livingston (L1) LIGO detectors is divided into $N = 121$ segments of duration $\Delta T = 25$ hours.

The phase evolution of the expected signal at the detector can be written as [28]

$$\begin{aligned} \phi(t) \approx & \phi_0 + 2\pi \sum_{k=0}^s \frac{f^{(k)}(t_0)(t-t_0)^{k+1}}{(k+1)!} \\ & + 2\pi \frac{\vec{r}(t)}{c} \vec{n} \sum_{k=0}^s \frac{f^{(k)}(t_0)(t-t_0)^k}{k!}, \end{aligned} \quad (8.1)$$

where α, δ are the standard equatorial coordinates, i.e., right-ascension and declination, ϕ_0 is the initial phase, $f^{(k)} \equiv \frac{d^k f}{dt^k}$ represent the time derivatives of the signal frequency f at the solar system barycenter (SSB) at reference time t_0 , c is the speed of light, $\vec{r}(t)$ is the vector pointing from the SSB to the detector and $\vec{n} \equiv (\cos \alpha \cos \delta, \sin \alpha \cos \delta, \sin \delta)$ is the unit vector pointing from the SSB to the gravitational-wave source. The $\mathcal{F}$ -statistic is the standard coherent technique to extract the CW signals from the detector noisy data. This statistic is the result of a matched filtering, consisting in the correlation of the data with a template represented by $\lambda \equiv \{\alpha, \delta, f, \dot{f}\}$, where the intrinsic amplitude h_0 , the inclination angle with respect to the line of sight ι , and the polarization angle ψ , as well as the initial phase ϕ_0 have been analytically maximized. In a fully coherent grid-based $\mathcal{F}$ -statistic search the number of templates increases with a high power as a function of the observation time, therefore such searches are not suitable for wide parameter-space all-sky surveys. However, the reduction of the coherent baseline in

a semicoherent search [64, 88] makes these techniques computationally feasible in a distributed computing environment like Einstein@Home and (usually) more sensitive at fixed computing cost [98].

The templates used to cover the parameter space are placed using the notion of mismatch [52, 51], which is the fractional loss of expected squared signal-to-noise ratio (SNR^2) ρ^2 at a random parameter space point λ compared to the true signal location λ_s . In the presence of noise, the mismatch is accessible through the fractional loss of actually measured SNR^2 , namely

$$\check{\mu} \equiv \frac{2\mathcal{F}(\lambda_s) - 2\mathcal{F}(\lambda)}{2\mathcal{F}(\lambda_s) - 4}. \quad (8.2)$$

To quadratic order the mismatch can also be computed using a symmetric positive-definite matrix g_{ij} , referred to as the parameter-space metric, namely

$$\mu^* \equiv g_{ij}(\lambda_s) \Delta\lambda^i \Delta\lambda^j, \quad (8.3)$$

where $\Delta\lambda^i = \lambda^i - \lambda_s^i$ and the summation convention over repeated indices apply. We use Eq. (8.3) as a distance measure.

In the S5R5 analysis, the templates at frequency f are placed on a coarse grid constructed using the spacings [59]

$$d\theta_{\mathcal{F}} = \frac{\sqrt{3}c}{v_d f \Delta T}, \quad df = \frac{\sqrt{12}m}{\pi \Delta T}, \quad d\dot{f} = \frac{\sqrt{3.3}m}{\Delta T^2}, \quad (8.4)$$

where m is the maximal mismatch per direction and v_d is the Earth's rotation speed at the equator. Note that the spindown spacing has been estimated in a Monte Carlo study and is finer than the expression derived using the metric, namely $d\dot{f} = \sqrt{720m}/\pi \Delta T^2$. As a consequence the fine grid is constructed only with sky refinement and without spindown refinement. However this is due to some practical limitations of the Einstein@Home environment, e.g., memory footprint, otherwise the spindown spacing would be refined by a factor N .

The resolution of the fine sky grid at frequency f is given by [59]

$$d\theta_H = \frac{c df}{\wp f v_y}, \quad (8.5)$$

where $\wp$ is the pixel factor and v_y is the Earth's orbital velocity. With $\wp = 0.5$, $m = 0.3$ the sky refinement used in the S5R5 search yields $\mathcal{N}_{sky}^{\text{ref}} = (d\theta_{\mathcal{F}}/d\theta_H)^2 \approx 8444$ [59].

To every parameter space point of the search has been assigned a significance (or critical ratio) CR [59]:

$$\text{CR} = \frac{n_c - \bar{n}_c}{\sigma}, \quad (8.6)$$

where

$$n_c = \sum_{i=0}^{N-1} w_i n_i \quad (8.7)$$

is the Hough number count with w_i being the weight for a frequency f and particular sky position (α, δ) , $n_i = 1$ if the $\mathcal{F}$ -statistic crosses a threshold value equal to 5.2 otherwise $n_i = 0$, $\bar{n}_c$ is the expected value of n_c and σ the standard deviation. The candidates are ordered by their significance, however in addition the average per segment $\overline{2\mathcal{F}}$ value is computed.

8.2. Follow-up method

8.2.1. The two-stage follow-up

The two-stage follow-up procedure has been introduced in Chap. 7. In the first semicoherent refinement stage we use the semicoherent $\mathcal{F}$ -statistic to improve the maximum-likelihood estimator (MLE) and in the second fully coherent zoom stage we use the fully coherent $\mathcal{F}$ -statistic to test the consistency of the candidate with Gaussian noise or its agreement with the signal model. We distinguish three possible outcomes of the zoom stage:

- *Consistency with Gaussian noise (G)* - the fully coherent $\widetilde{2\mathcal{F}}$ value does not exceed a threshold, i.e.,

$$\widetilde{2\mathcal{F}} < G_{\text{th}}, \quad (8.8)$$

where G_{th} is properly chosen to corresponds to some (small) false-alarm probability p_{fA} in Gaussian noise. The single trial false-alarm probability for a given G_{th} threshold is $p_{\text{fA}}^1 = (1 + \mathcal{F})e^{-\mathcal{F}}$ (see, e.g., [114] for details). For example, a threshold $G_{\text{th}} = 90$ corresponds to a false-alarm probability of order 10^{-12} for a single template. Assuming $\mathcal{N}$ independent templates and $p_{\text{fA}} \ll 1$, the false-alarm is $p_{\text{fA}} \approx \mathcal{N}p_{\text{fA}}^1$.

- *Non-Gaussian origin ($\neg G$)* - the candidate is loud enough to be inconsistent with Gaussian noise at the false-alarm probability p_{fA} , i.e.,

$$\widetilde{2\mathcal{F}} \geq G_{\text{th}}. \quad (8.9)$$

- We define *signal recovery (S)* as a *subclass* of $\neg G$, namely a signal is considered recovered if for the final zoomed candidate the $\widetilde{2\mathcal{F}}$ -value is exceeding the Gaussian-noise threshold G_{th} and falling into a predicted signal interval

$$\widetilde{2\mathcal{F}}_{\text{th}}^{(S)} < \widetilde{2\mathcal{F}} < \widetilde{2\mathcal{F}}_{\text{max}}^{(S)} \quad (8.10)$$

where $\widetilde{2\mathcal{F}}_{\text{th}}^{(S)} \equiv \max\{G_{\text{th}}, \widetilde{2\mathcal{F}}_o - n_u \sigma_o\}$, and $\widetilde{2\mathcal{F}}_{\text{max}}^{(S)} \equiv \widetilde{2\mathcal{F}}_o + n_u \sigma_o$, with expectation

$$\widetilde{2\mathcal{F}}_o \approx 4 + N \overline{\rho^2}_{\text{MLE}} \quad (8.11)$$

and variance

$$\sigma_o^2 \approx 2(4 + 2N \overline{\rho^2}_{\text{MLE}}) ; \quad (8.12)$$

$\overline{\rho^2}_{\text{MLE}}$ is the average SNR^2 of the MLE of the semicoherent search and n_u determines the desired confidence level. For example $n_u = 6$ corresponds roughly to a confidence of $\sim 99.6\%$.

In both stages the parameter space is explored on a dynamically created mesh by a MADS-based algorithm. The input to the algorithm is a starting point λ_c , a search bounding box $\Delta\lambda$ around the candidate and a set of MADS parameters, namely $\{d\lambda, u_b, w_{\min}^+, w_{\max}^+, w^-\}$, where $d\lambda$ is the initial step, u_b is the mesh update basis, $w_{\min}^+$ and $w_{\max}^+$ are the mesh-coarsening exponents and w^- denotes the mesh-refining exponent; for details we refer to Chap. 7.2.1. However, in this chapter we stress an important difference in the usage of the bounding box $\Delta\lambda_R$ with respect to the treatment in Chap. 7. In the previous chapter the search in the refinement stage is effectively done in a parameter space defined by the semicoherent metric ellipsoid. Here the bounding box of the refinement stage is fixed to some empirically found size and the optimization is constrained by the full size of the box. This is required by the setup of the original Hough-transform search. In fact, a pre-refinement Hough-transform semicoherent refinement stage yields even better overall performance of the follow-up procedure.

8.2.2. A pre-refinement with the semicoherent Hough-transform search

Since we aim to explore a very large parameter space at a reasonable computing cost, the original semicoherent Hough-transform search discussed in [59] makes use of mismatch, defined by Eq. (8.3), that is large enough, such that $\check{\mu} \ll \mu^*$. Given that we have only a handful of candidates to follow-up, and we can therefore afford the computational cost, it is reasonable to introduce a pre-refinement, e.g., by using the semicoherent Hough-transform search, in order to improve the candidate for the refinement stage. We refer to this stage as the semicoherent Hough refinement (HS_{ref}). The refinement is achieved by increasing the original spindown resolution $d\dot{f}$ by the number of the segments used in the search, i.e., $N = 121$. In addition, we also double the pixel factor found in Eq. (8.5), thus increasing the sky grid refinement. The loudest candidate selected from this initial step provides the starting point for the grid-less optimization with the MADS-based algorithm.

8.2.3. The number of search dimensions

In this subsection we briefly discuss the number of search dimensions, i.e., the spindown order to be included in the follow-up procedure. For a grid based search, the minimal required spindown order to be included in the search maximizes the number of templates [64, 88], i.e., the parameter-space extent in a given direction must be thicker than the metric template extent in that direction as discussed in Chap. 4. On the other hand, a grid-less search can be considered as a grid-based search with a template bank constructed by using an unknown, in general non-constant mismatch, for any single

template. Clearly the maximization of the number of templates criteria cannot be applied. How to adopt the parameter-space extent criteria in this case and decide, which spindown-order to use, requires some more thought. However, as a rule of thumb, the longer the time spanned by the data, the better to include higher-order spindown in the follow-up. In this chapter we restrict ourselves to the follow-up procedure with searches including only the first-order spindown. Nevertheless, in Sec. 8.3.2 we perform a Monte Carlo of a zoom stage searching for signals with non-zero second-order spindown.

8.3. Monte Carlo studies

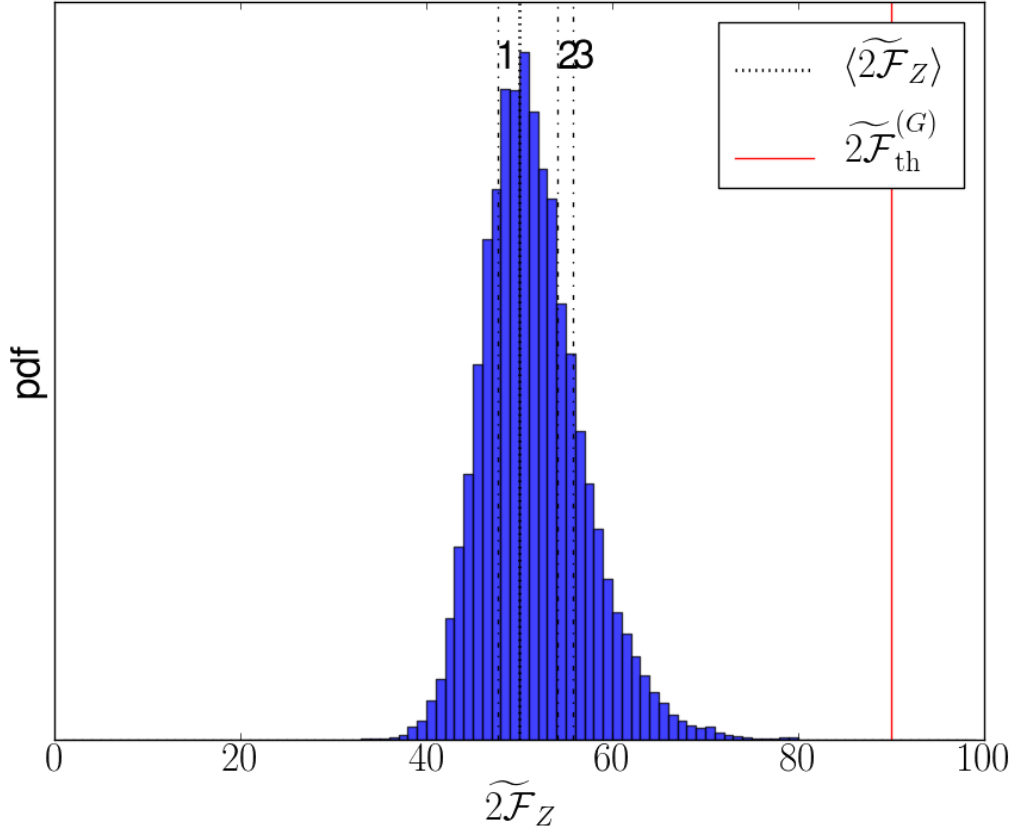


Figure 8.1.: $\widetilde{2\mathcal{F}}_Z$ distribution after the fully coherent 4-D $\{\alpha, \delta, f, \dot{f}\}$ zoom stage of 15000 searches in pure Gaussian noise data without injected signals. The maximal found value is $\widetilde{2\mathcal{F}}_Z^{max} = 79.39$. The mean value is $\langle \widetilde{2\mathcal{F}}_Z \rangle = 51.53$ plotted with a dotted line. The candidates at roughly 434, 677 and 984 Hz are also plotted and labeled with 1, 2, 3.

We test the proposed follow-up pipeline in an end-to-end Monte Carlo study using

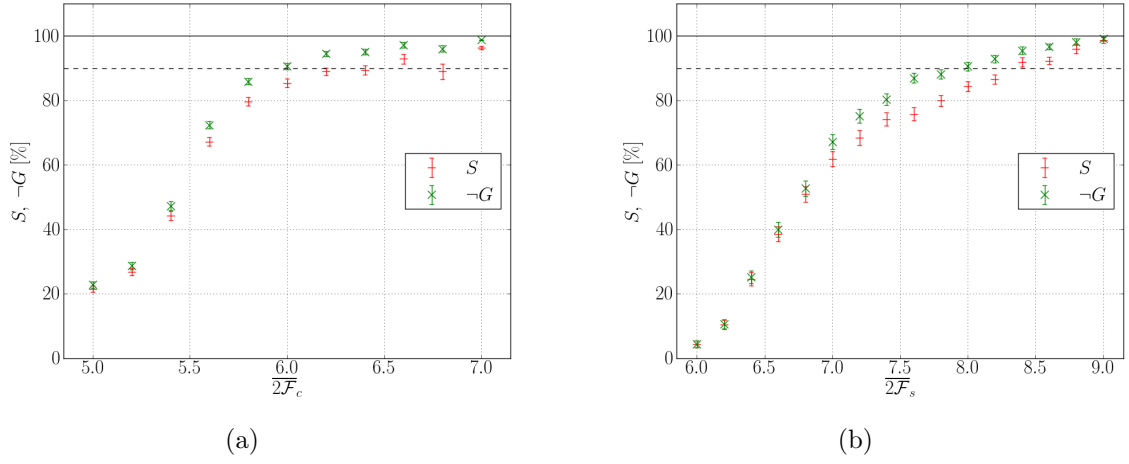


Figure 8.2.: Monte Carlo study of the efficiency of the follow-up pipeline with initial hierarchical search Hough refinement. (a) Percentage of the injected signals classified as recovered ($-$ S) and with a non-Gaussian origin ($\times$ $\neg G$) as function of the average $2\mathcal{F}$ value of the candidate after the original semicoherent Hough-transform search $2\overline{\mathcal{F}}_c$. (b) Percentage of the injected signals classified as recovered ($-$ S) and with a non-Gaussian origin ($\times$ $\neg G$) as function of the average injected $2\mathcal{F}$ value, $2\overline{\mathcal{F}}_s$. In (a) and (b) the error bars are computed by using a Jackknife estimator.

the LALSUITE [105] software package. In particular we use the LALAPPS applications `Makefakedata_v4` to generate Gaussian noise and inject CW signal, `HierarchicalSearch` for the semicoherent Hough-transform search, `FStatSCNomad` for the semicoherent $\mathcal{F}$ -statistic optimization with MADS and `FStatFCNomad` for the fully coherent $\mathcal{F}$ -statistic MADS optimization, where for the MADS algorithm we use the reference implementation NOMAD [101]. Whenever we need to compute the fully coherent or semicoherent metric, we use the application `FstatMetric_v2`.

We apply the follow-up chain to 15000 different noise realizations with and without injected signal. The Gaussian noise realizations are generated with the `MakeFakedata_v4` application using the same timestamps of the SFTs¹ used in the original Einstein@Home search with level of the noise per detector set to $1.99 \times 10^{-23} \text{ Hz}^{-1/2}$. The signal parameters are uniformly drawn in the ranges $\cos \iota \in [-1, 1]$, $\psi \in [-\pi/4, \pi/4]$, $\phi_0 \in [0, 2\pi]$ and $f \in [185, 186] \text{ Hz}$, except the sky position which is drawn isotropically on the sky. The frequency range has been chosen in the most sensitive region of detectors. The spindown value is randomly chosen in the range $\dot{f} \in (-\frac{f_0}{\tau_0}, 0.1\frac{f_0}{\tau_0})$ with minimal spindown age $\tau_0 = 800 \text{ y}$ at $f_0 = 50 \text{ Hz}$. The signal amplitude is high enough such that the SNR^2 in the point of injection is uniformly distributed in the range $\bar{\rho}^2 \in [0, 6]$.

We begin the end-to-end validation with a simulation stage of the original S5R5 Einstein@Home search by using the original search setup, i.e., the same frequency and spindown grid spacings given by Eq. (8.4). The S5R5 search has been partitioned in independent computing tasks, referred to as workunits (WUs). For a detailed discussion of the workunit design see Sec. III C in [59]. To save computing power, we do not rerun an entire WU in this simulation stage, but we center a search grid around a random point in the vicinity of the injected signal, searching 10 frequency bins in total. The sky-grid is constructed by extraction of the 16 closest points to the candidate from the original sky-grid file. However, this reduced parameter-space size is still sufficiently large to be possible to select candidates due to the noise, if the signal is weak as it might happen in a real search, and not artificially select a point close to the true signal location.

In the pre-refinement semicoherent Hough search we center the coarse grid from the simulation around the most significant candidate found so far. However for this pre-refinement search we refine the spindown spacing of the fine grid by a factor $N = 121$ and set the pixel factor in Eq. (8.6) equal to 1. In this way we improve the starting point for the MADS-based refinement.

The semicoherent $\mathcal{F}$ -statistic refinement with the MADS-based algorithm uses a fixed search box around the most significant candidate found in the Hough refinement search. The parameter space is restricted in a box with size given by

$$\begin{aligned} \Delta\alpha &= 0.4 \text{ rad}, \quad \Delta\delta = 0.4 \text{ rad} \\ \Delta f &= 1 \times 10^{-4} \text{ Hz}, \quad \Delta\dot{f} = 1 \times 10^{-9} \text{ Hz/s}. \end{aligned} \tag{8.13}$$

¹SFT is the acronym used for Short time baseline Fourier Transform of the calibrated detector strain data. The duration of the SFTs is typically 1800 seconds. SFTs are used as input to many CW searches such as the semicoherent Hough-transform search, as well as the fully coherent follow-up.

These values have been found empirically to yield good results. The loudest point selected from this step is used as a starting point for the fully coherent $\mathcal{F}$ -statistic zoom search.

The zoom search is constrained by the Fisher ellipse as described in the previous chapter, Eq. (7.25), where we use a confidence corresponding to 24 standard deviations. This large number is chosen to not falsely miss weak signals, for which the Fisher matrix may be a poor predictor.

The algorithm parameters for the MADS-based refinement and zoom stage are summarized in Table 8.1. These parameters have been found to yield good results in Monte Carlo studies.

stage	w^-	$w_{\min}^+$	$w_{\max}^+$	u_b	p
R	2	1	20	2	20000
Z	2	1	50	1.2	20000

Table 8.1.: Follow-up algorithm parameters for the refinement and zoom stage.

8.3.1. Efficiency of the follow-up pipeline

We first run the end-to-end validation with semicoherent refinement using the Gaussian noise data without any injected signals. This is required to ensure the applicability of the threshold $G_{\text{th}} = 90$ used to consider a candidate as conform with the Gaussian noise hypothesis. The distribution of the $2\mathcal{F}_Z$ values is plotted in Fig. 8.1, where we also show the outcome of the follow-up for 3 loud S5R5 candidates (details are given in Sec. 8.5). The maximal value found is $\widetilde{2\mathcal{F}}_Z^{\max} = 79.39$, which is well below $\neg G = 90$.

In Fig. 8.2(a) we plot the percentage of the injected signals classified as recovered and with a non-Gaussian origin as a function of the average $2\mathcal{F}$ -value of the candidate after the simulation stage. We are able to separate 90 % of the candidates from Gaussian noise starting with $\overline{2\mathcal{F}}_c \approx 6.0$ and we recover 90 % of the signals for candidates with $\overline{2\mathcal{F}}_c \approx 6.2$.

8.3.2. Impact of second-order spindown on the first-order spindown follow-up

Before proceeding with the follow-up studies using *S5R5* candidates and data, we test the response of the fully coherent stage of the first-order follow-up pipeline when searching for a signal with non-zero second-order spindown. To do this, we simulate 1000 Gaussian noise realizations, and for each of them, we draw random signal parameters, as previously discussed, but with fixed frequency $f = 50.25$ Hz and first-order spindown value $\dot{f} \approx -2 \times 10^{-9}$. For each trial we vary the second-order spindown value uniformly distributing

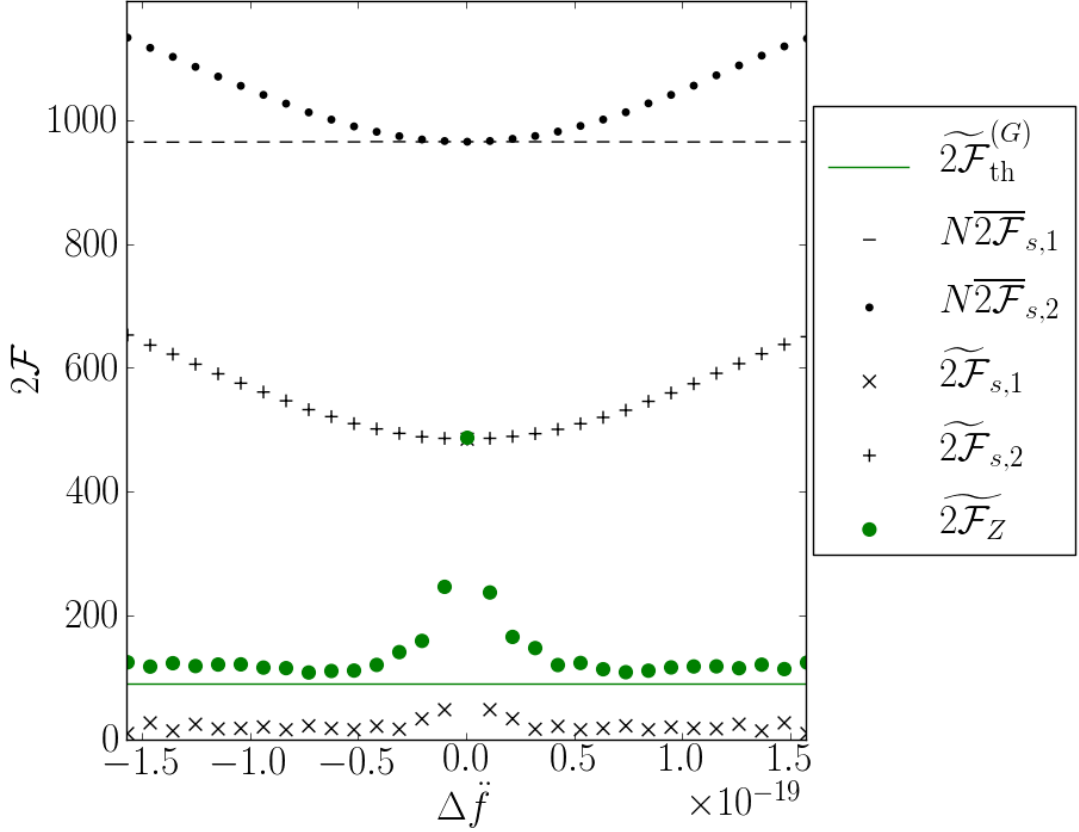


Figure 8.3.: Monte Carlo study of the efficiency of a fully-coherent MADS-based $\mathcal{F}$ -statistic search over 4D space, i.e., $\{\alpha, \delta, f, \dot{f}\}$, when a signal with second-order spindown $\ddot{f} \in [-2f_0/\tau_0^2, 2f_0/\tau_0^2]$ is present in the data. We plot the averaged over 1000 trials $N\overline{2\mathcal{F}}_{s,1}$, $N\overline{2\mathcal{F}}_{s,2}$, $\widetilde{2\mathcal{F}}_{s,1}$, $\widetilde{2\mathcal{F}}_{s,2}$ and $\widetilde{2\mathcal{F}}_Z$ as function of the deviation $\Delta\ddot{f}$ of the second-order spindown from zero. Note that for each trial $\overline{2\mathcal{F}}_{s,1} \approx 8.0$. In 99.9 % of the trials $\widetilde{2\mathcal{F}}_Z > G_{\text{th}}$.

it in the range

$$\ddot{f} \in \left[-\frac{2f_0}{\tau_0^2}, \frac{2f_0}{\tau_0^2} \right], \quad (8.14)$$

and add the signal to the simulated Gaussian noise data.

Let the signal template be $\lambda_{s,2} \equiv \lambda_s = (\alpha_s, \delta_s, f_s, \dot{f}_s, \ddot{f}_s)$, with a $\ddot{f}_s$ value from the range given in [8.14], and let the truncated signal template be $\lambda_{s,1} = (\alpha_s, \delta_s, f_s, \dot{f}_s, 0)$. We denote the $2\mathcal{F}$ statistic computed in the point $\lambda_{s,1}$, resp. $\lambda_{s,2}$ with $2\mathcal{F}_{s,1}$, resp. $2\mathcal{F}_{s,2}$. With this notation, the signal amplitude has been chosen such that independently of the deviation $\Delta\ddot{f}$ of the second-order spindown from zero, we measure $\overline{2\mathcal{F}}_{s,1}$ in the range $\overline{2\mathcal{F}}_{s,1} \in (7.9, 8.1)$. This particular range for $\overline{2\mathcal{F}}_{s,1}$ has been chosen, because it has been found sufficient to cross the threshold $G_{\text{th}} = 90$ for 90 % in the trials of the end-to-end pipeline validation from the previous subsection; see Fig. 8.2(b) where we plot the percentage of the injected signals, classified as recovered or at least with a non-Gaussian origin, as function of the averaged injected $\overline{2\mathcal{F}}_s$. First, we measure $\overline{\mathcal{F}}_{s,2}$, $\widetilde{2\mathcal{F}}_{s,1}$, as well as $2\mathcal{F}_{s,2}$. Then we perform a zoom stage using the parameters given in Table 8.1, with starting point $\lambda_{s,1}$. The results of this Monte Carlo study are plotted in Fig. 8.3. For convinience, we plot $N\overline{2\mathcal{F}}_{s,1}$ and $N\overline{2\mathcal{F}}_{s,2}$ instead of $\overline{2\mathcal{F}}_{s,1}$ and $\overline{2\mathcal{F}}_{s,2}$. Note that for $|\Delta\ddot{f}| > 0$ the ratio $\overline{2\mathcal{F}}_{s,2}/\overline{2\mathcal{F}}_{s,1}$ is much smaller compared to the ratio $\widetilde{2\mathcal{F}}_{s,2}/\widetilde{2\mathcal{F}}_{s,1}$. This is expected due to the short duration of the segments used to compute the semicoherent statistic, compared to the total observation time used to compute the fully coherent statistic. However, by applying the zoom stage, even if we do not include the second-order spindown in the search, we are able to find a parameter-space point louder than the threshold G_{th} in 99.9% of the cases for the given second-order spindown range. While this is promissing, a more conclusive result requires additional work in order to validate the full pipeline when searching signals with second-order spindown.

8.4. Recovery of hardware injections

The hardware injections are simulated signals, physically added into the control system of the interferometer to produce a detector response similar to what should be generated if a gravitational wave were present. The aim of such injections is to test and evaluate search pipelines and analysis codes. The S5R5 Einstein@Home search [59] identified 3 fake pulsars referred to as Pulsar 2, 3 and 5. In this section we follow-up the candidates associated with these hardware injections, since all of them produced more than one significant candidate to follow-up. Instead to further investigate all these candidate points, we apply a clustering algorithm to reduce their number and thus follow-up the most interesting ones.

The clustering algorithm is rather simple and described in the following. For each hardware injection starting with the loudest candidate, we remove all neighboring candidates falling in the box defined by the search bands used for the refinement stage as given in Eq. (8.13). We expect that these candidates would have a similar maximum,

as they are most likely due to the same signal. We then repeat the procedure with the remaining candidates until there are no more candidates left sharing a common semicoherent search box.

There are for instance 88 parameter space points associated with Pulsar 2 injected at ~ 575 Hz. After the clustering procedure, the number of candidates to follow-up is reduced to 16. For Pulsar 3, injected at ~ 108 Hz, the number of parameter space points to follow-up shrinks from 80 to 9. Only for Pulsar 5, injected at ~ 52 Hz, we are not able to reduce the number of candidates, since there are only 2 candidate points to follow-up, but they belong to different search boxes.

In Table 8.2 we summarize the follow-up results of the loudest candidates associated with each pulsar. Note that the hardware injections were only partially active during the total observation time (for details see Sec. VI of [59]). Thus, in order to consider a signal recovered we rely on the measured fully coherent $2\mathcal{F}$ -value in the point of injection $2\mathcal{F}_s$. Based on this, we see that all three candidates are recovered. Pulsar 3 is the best recovered hardware injection in terms of parameter estimation with a mismatch after the final step of $\mu_Z^* = 1.44 \times 10^{-3}$. Pulsar 2 has been recovered fairly well with a mismatch $\mu_Z^* = 2.41 \times 10^{-3}$, though the spindown value is found with opposite sign. Finally, the mismatch for Pulsar 5 is $\mu_Z^* = 6.72 \times 10^{-3}$. In Table 8.2 we also give the offset in frequency Δf_Z and spindown $\Delta \dot{f}_Z$ of the loudest point compared to the true signal location and the angular distance

$$\Delta\gamma_Z = \arccos(\vec{n}_c \vec{n}_s) . \quad (8.15)$$

8.5. Follow-up of loud S5R5 Einstein@Home candidates

The post-processing and further investigation of candidates from the S5R5 Einstein@Home run provides 9 interesting parameter space points for the fully coherent follow-up search at roughly 80, 96, 144, 434, 677, 932, 984, 1030 and 1141 Hz. We investigate these loud candidates using the pipeline with a hierarchical search refinement and algorithm parameters from the Monte Carlo. The results are summarized in Table 8.3. None of the candidates can be considered as recovered signal in the sense of (8.12). However we can separate the candidates into 2 groups. First, the candidates at roughly 80, 96, 144, 932, 1030 and 1141 Hz are discarded, since they fail the multi-detector versus single-detector $\mathcal{F}$ -statistic consistency test after the semicoherent MADS-based refinement stage. The $\mathcal{F}$ -statistic consistency test requires, that the combined $2\mathcal{F}$ statistic from N_{det} detectors is higher than any individual one (for details see Sec. IV C in [59]). On the other hand the candidates at roughly 434, 677, and 984 Hz survive the consistency check. In Fig. 8.1 we plot them on top of the $2\mathcal{F}_Z$ distribution of Gaussian noise. We see that the final statistics for these candidates is consistent with the expectation from Gaussian noise.

Row	Pulsar	2	3	5
1	f [Hz]	575.163557631402	108.857159397497	52.8083243593611
2	α [rad]	3.75692884	3.113188712	5.281831296
3	δ [rad]	0.060108958	-0.583578803	-1.463269033
4	$\dot{f}$ [Hz/s]	-1.37×10^{-13}	-1.46×10^{-17}	-4.03×10^{-18}
5	$\overline{2\mathcal{F}}_{\text{HS}}$	27.95	339.41	6.34
6	$\overline{2\mathcal{F}}_R$	100.54	1137.97	12.01
7	$\overline{2\mathcal{F}}_R^{\text{H1}}$	51.17	641.77	8.24
8	$\overline{2\mathcal{F}}_R^{\text{L1}}$	54.54	510.54	8.01
9	μ_R^*	4.01×10^{-4}	5.18×10^{-4}	4.88×10^{-3}
10	f_Z [Hz]	575.1635576321457	108.8571593975228	52.80832435480315
11	α_Z [rad]	3.756928870833068	3.113189005400489	5.281811486307014
12	δ_Z [rad]	0.06010925352403444	-0.5835788464819518	-1.463265695745091
13	$\dot{f}_Z$ [Hz/s]	-1.37×10^{-13}	3.30×10^{-16}	1.85×10^{-15}
14	$\widetilde{2\mathcal{F}}_Z$	7399.52	87097.9	678.06
15	$\widetilde{2\mathcal{F}}_Z^{\text{H1}}$	3519.49	47572.90	350.00
16	$\widetilde{2\mathcal{F}}_Z^{\text{L1}}$	3896.15	39557.50	332.64
17	$2\mathcal{F}_s$	7377.94	86968.70	677.86
18	μ_Z^*	2.41×10^{-3}	1.44×10^{-3}	6.72×10^{-3}
19	Δf_Z [Hz]	7.44×10^{-10}	2.58×10^{-11}	-4.55×10^{-9}
20	$\Delta \dot{f}_Z$ [Hz/s]	-4.33×10^{-16}	3.44×10^{-16}	1.85×10^{-15}
21	$\Delta \gamma_Z$ [rad]	2.97×10^{-7}	2.48×10^{-7}	3.96×10^{-6}

Table 8.2.: Loudest candidates associated with the hardware injection recovery. Values obtained using only H1 or L1 detector data are denoted with superscript H1, resp. L1. In rows 1 to 4 are given the true parameters of the injections. Row 5 is the average $2\mathcal{F}$ value after the semi-coherent Hough-transform search. In rows 6 to 8 are given the average $2\mathcal{F}$ values after the refinement stage. In row 9 is given the metric mismatch after the refinement. Rows 10 to 13 show the parameters of the candidate after the zoom stage. The $2\mathcal{F}$ values after the zoom stage are given in rows 14 to 16. In row 17 is given the fully coherent $2\mathcal{F}$ value of the injection. In row 18 we show the mismatch to the signal after the zoom stage. In rows 19 to 21 are given the offsets to the signal after the zoom stage. Due to the signal strength we are able to estimate the parameters of the signal much better than one would expect from the intrinsic resolution of the parameter space.

f [Hz]	$\overline{2\mathcal{F}}_R$	$\overline{2\mathcal{F}}_R^H$	$\overline{2\mathcal{F}}_R^L$	$\widetilde{2\mathcal{F}}_Z$	$\widetilde{2\mathcal{F}}_Z^H$	$\widetilde{2\mathcal{F}}_Z^L$	$\widetilde{2\mathcal{F}}_o$	$\widetilde{2\mathcal{F}}_{\text{th}}^{(S)}$	status
80	7.44	4.75	8.13	61.95	26.08	37.30	420.88	175.28	fcc
96	9.13	4.47	13.18	76.52	26.50	56.77	625.25	325.68	fcc
144	11.17	4.57	14.02	79.83	26.93	55.11	871.84	517.93	fcc
434	5.54	5.42	4.58	47.75	30.19	22.00	190.17	90.00	cgn
677	6.45	5.40	5.23	54.11	44.51	14.09	300.97	93.48	cgn
932	7.60	8.04	4.25	73.40	66.62	16.94	439.25	188.37	fcc
984	6.50	4.82	5.49	55.81	36.30	20.90	305.94	96.73	cgn
1030	7.44	8.30	4.51	79.31	71.63	15.20	420.39	174.94	fcc
1141	8.46	10.13	4.23	75.33	97.26	12.38	543.71	264.42	fcc

Table 8.3.: Summary of the follow-up results for the nine candidates from the S5R5 search. The last column gives the reason, why the candidate has been discarded, where *fcc* means “failed detector consistency check”, and *cgn* means “consistent with Gaussian noise”. Only three of these candidates pass the detector consistency check after the refinement stage, however they are consistent with Gaussian noise after the zoom stage.

8.6. Discussion

We extended the two-stage follow-up procedure for the fully coherent follow-up of CW candidates from semicoherent searches with a pre-refinement semicoherent Hough-transform search in order to investigate loud candidates from the S5R5 Einstein@Home search for periodic gravitational waves in the LIGO S5 data. We studied the efficiency of the proposed follow-up pipeline on simulated signals in simulated Gaussian noise data, and we discussed the impact of a possible deviation of the second-order spindown from zero. We followed-up 27 candidates associated with 3 hardware injections and successfully recovered all 3 signals. We followed-up 9 loud candidates of unknown origin. None of these candidates could be confirmed to follow the signal model. After the refinement stage 6 of the candidates have been discarded failing the detector consistency check. The remaining 3 candidates are consistent with Gaussian noise after the fully coherent zoom stage, thus no significant CW candidate have been found.

This is the first application of the systematic follow-up procedure on real gravitational-wave detector data. While additional work is required to include higher order spindown in the final fully coherent stage, we expect that the presented strategy will become one of the common tools applied in the final stages of wide parameter-space searches for CWs in the upcoming Advanced detectors era.

Bibliography

- [1] B. Schutz, *A First Course in General Relativity*. Cambridge University Press, 2nd ed., June 2009.
- [2] R. M. Wald, *General Relativity*. University Of Chicago Press, 1984.
- [3] M. Maggiore, *Gravitational Waves: Volume 1: Theory and Experiments*. Gravitational Waves, Oxford University Press, USA, 2007.
- [4] P. Jaranowski and A. Krolak, *Analysis of Gravitational-Wave Data*. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 2009.
- [5] W. Anderson and J. D. E. Creighton, *Gravitational-Wave Physics and Astronomy: An Introduction to Theory, Experiment and Data Analysis*. Wiley-VCH, Nov. 2011.
- [6] D. Sivia, *Data Analysis: A Bayesian Tutorial*. Oxford Science Publications, Oxford University Press, USA, 1996.
- [7] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (v. 1)*. Prentice Hall, 1 ed., Jan. 1993.
- [8] B. Sathyaprakash and B. Schutz, “Physics, Astrophysics and Cosmology with Gravitational Waves,” *Living Rev.Rel.*, vol. 12, p. 2, 2009.
- [9] B. F. Schutz and F. Ricci, “Gravitational Waves, Sources, and Detectors,” 2010.
- [10] E. E. Flanagan and S. A. Hughes, “The Basics of gravitational wave theory,” *New J.Phys.*, vol. 7, p. 204, 2005.
- [11] R. Prix, “Gravitational waves from spinning neutron stars,” in *Neutron Stars and Pulsars* (W. Becker, ed.), vol. 357 of *Astrophysics and Space Science Library*, p. 651, Springer Berlin Heidelberg, 2009.
- [12] S. Luke, *Essentials of Metaheuristics*. Lulu, 2009. Available for free at <http://cs.gmu.edu/~sean/book/metaheuristics/>.
- [13] T. G. Kolda, R. M. Lewis, and V. Torczon, “Optimization by direct search: New perspectives on some classical and modern methods,” *SIAM Review*, vol. 45, pp. 385–482, 2003.

- [14] D. Kennefick, “Controversies in the history of the radiation reaction problem in general relativity,” 1997.
- [15] J. Weisberg, D. Nice, and J. Taylor, “Timing Measurements of the Relativistic Binary Pulsar PSR B1913+16,” *Astrophys.J.*, vol. 722, pp. 1030–1034, 2010.
- [16] J. Weber, “Evidence for discovery of gravitational radiation,” *Phys. Rev. Lett.*, vol. 22, pp. 1320–1324, Jun 1969.
- [17] O. D. Aguiar, “The Past, Present and Future of the Resonant-Mass Gravitational Wave Detectors,” *Res.Astron.Astrophys.*, vol. 11, pp. 1–42, 2011.
- [18] M. Cerdonio, L. Conti, J. Lobo, A. Ortolan, and J. Zendri, “Wideband dual sphere detector of gravitational waves,” *Phys.Rev.Lett.*, vol. 87, p. 031101, 2001.
- [19] M. Bonaldi, M. Cerdonio, L. Conti, P. Falferi, P. Leaci, *et al.*, “Principles of wide bandwidth acoustic detectors and the single-mass DUAL detector,” *Phys.Rev.*, vol. D74, p. 022003, 2006.
- [20] P. Leaci, T. D. Canton, A. Ortolan, and G. A. Prodi, “Antenna pattern of DUAL detectors of gravitational waves and its exploitation in a network of advanced interferometers,” *Classical and Quantum Gravity*, vol. 25, no. 19, p. 195018, 2008.
- [21] B. Abbott *et al.*, “Detector description and performance for the first coincidence observations between LIGO and GEO,” *Nucl.Instrum.Meth.*, vol. A517, pp. 154–179, 2004.
- [22] M. Pitkin, S. Reid, , and J. Hough, “Gravitational wave detection by interferometry (ground and space),” *Living Reviews in Relativity*, vol. 14, no. 5, 2011.
- [23] S. A. Hughes and K. S. Thorne, “Seismic gravity-gradient noise in interferometric gravitational-wave detectors,” *Phys. Rev. D*, vol. 58, p. 122002, Nov 1998.
- [24] “A Gravitational wave observatory operating beyond the quantum shot-noise limit: Squeezed light in application,” *Nature Phys.*, vol. 7, pp. 962–965, 2011.
- [25] D. Shoemaker, “Advanced LIGO reference design,” tech. rep., 2011.
- [26] K. Somiya, “Detector configuration of KAGRA: The Japanese cryogenic gravitational-wave detector,” *Class.Quant.Grav.*, vol. 29, p. 124007, 2012.
- [27] H. L. et al., “Einstein gravitational wave Telescope conceptual design study,” tech. rep., 2011.
- [28] P. Jaranowski, A. Krolak, and B. F. Schutz, “Data analysis of gravitational-wave signals from spinning neutron stars. I: The signal and its detection,” *Phys. Rev. D.*, vol. 58, p. 063001, 1998.

- [29] . J. Abadie *et al.*, “All-sky search for gravitational-wave bursts in the second joint LIGO-Virgo run,” 2012.
- [30] M. Briggs *et al.*, “Search for gravitational waves associated with gamma-ray bursts during LIGO science run 6 and Virgo science runs 2 and 3,” 2012.
- [31] S. Adrian-Martinez *et al.*, “A First Search for coincident Gravitational Waves and High Energy Neutrinos using LIGO, Virgo and ANTARES data from 2007,” 2012.
- [32] W. Becker, *Neutron Stars and Pulsars*. Astrophysics and Space Science Library, Springer, 2008.
- [33] G. Ushomirsky, C. Cutler, and L. Bildsten, “Deformations of accreting neutron star crusts and gravitational wave emission,” *Mon.Not.Roy.Astron.Soc.*, vol. 319, p. 902, 2000.
- [34] A. Worley, P. G. Krastev, and B.-A. Li, “Nuclear constraints on the moments of inertia of neutron stars,” *The Astrophysical Journal*, vol. 685, no. 1, p. 390, 2008.
- [35] N. Stergioulas, “Rotating stars in relativity,” *Living Reviews in Relativity*, vol. 6, no. 3, 2003.
- [36] B. J. Owen, L. Lindblom, C. Cutler, B. F. Schutz, A. Vecchio, *et al.*, “Gravitational waves from hot young rapidly rotating neutron stars,” *Phys.Rev.*, vol. D58, p. 084020, 1998.
- [37] C. Van Den Broeck, “The Gravitational wave spectrum of non-axisymmetric, freely precessing neutron stars,” *Class.Quant.Grav.*, vol. 22, pp. 1825–1840, 2005.
- [38] W. Baade and F. Zwicky, “Remarks on super-novae and cosmic rays,” *Phys. Rev.*, vol. 46, pp. 76–77, Jul 1934.
- [39] A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, and R. A. Collins, “Observation of a Rapidly Pulsating Radio Source,” *Nature*, vol. 217, p. 709713, Feb. 1968.
- [40] P. Haensel, A. Potekhin, and D. Yakovlev, *Neutron Stars 1: Equation of State and Structure*. Astrophysics and Space Science Library, Springer, 2010.
- [41] R. Prix, “Search for continuous gravitational waves: metric of the multi-detector $\mathcal{F}$ -statistic,” *Phys. Rev. D.*, vol. 75, p. 023004, 2007.
- [42] J. H. Taylor and J. M. Weisberg, “Further experimental tests of relativistic gravity using the binary pulsar PSR 1913 + 16,” *ApJ*, vol. 345, pp. 434–450, Oct. 1989.
- [43] B. Krishnan *et al.*, “The Hough transform search for continuous gravitational waves,” *Phys. Rev. D.*, vol. 70, p. 082001, 2004.

- [44] R. Prix, “The $\mathcal{F}$ -statistic and its implementation in ComputeFStatistic_v2,” tech. rep., 2010.
- [45] J. Whelan, “Visualization of Antenna Pattern Factors via Projected Detector Tensors,” tech. rep., 2012.
- [46] C. Cutler and B. F. Schutz, “The generalized $\mathcal{F}$ -statistic: multiple detectors and multiple gravitational wave pulsars,” *Phys. Rev. D.*, vol. 72, p. 063006, 2005.
- [47] R. J. Dupuis and G. Woan, “Bayesian estimation of pulsar parameters from gravitational wave data,” *Phys. Rev.*, vol. D72, p. 102002, 2005.
- [48] L. S. Finn, “Detection, measurement, and gravitational radiation,” *Phys. Rev. D.*, vol. 46, pp. 5236–5249, Dec 1992.
- [49] C. Kittel, *Elementary Statistical Physics*. John Wiley and Sons, Inc., USA, 1967.
- [50] B. F. Schutz and M. A. Papa, “End-to-end algorithm for hierarchical area searches for long duration GW sources for GEO-600,” 1999.
- [51] B. J. Owen, “Search templates for gravitational waves from inspiraling binaries: Choice of template spacing,” *Phys. Rev. D.*, vol. 53, p. 6749, June 1996.
- [52] R. Balasubramanian, B. S. Sathyaprakash, and S. V. Dhurandhar, “Gravitational waves from coalescing binaries: Detection strategies and monte carlo estimation of parameters,” *Phys. Rev. D.*, vol. 53, p. 3033, 1996.
- [53] P. R. Brady, T. Creighton, C. Cutler, and B. F. Schutz, “Searching for periodic sources with LIGO,” *Phys. Rev. D.*, vol. 57, p. 2101, 1998.
- [54] R. Prix, “Template-based searches for gravitational waves: efficient lattice covering of flat parameter spaces,” *Class. Quant. Grav.*, vol. 24, p. S481, July 2007.
- [55] C. Messenger, R. Prix, and M. A. Papa, “Random template banks and relaxed lattice coverings,” *Phys. Rev. D.*, vol. 79, p. 104017, May 2009.
- [56] H. J. Pletsch, “Parameter-space metric of semicoherent searches for continuous gravitational waves,” *Phys. Rev. D.*, vol. 82, p. 042002, Aug. 2010.
- [57] “Einstein@home project.”
- [58] B. Abbott *et al.*, “Einstein@home search for periodic gravitational waves in ligo s4 data,” *Phys. Rev. D.*, vol. 79, p. 022001, Jan. 2009.
- [59] J. Aasi *et al.*, “Einstein@home all-sky search for periodic gravitational waves in ligo s5 data,” *Phys. Rev. D.*, vol. 87, p. 042001, Feb 2013.

- [60] J. P. Ostriker and J. E. Gunn, “On the Nature of Pulsars. I. Theory,” *ApJ*, vol. 157, p. 1395, Sept. 1969.
- [61] Y. Yue, R. Xu, and W. Zhu, “What can the braking indices tell us about the nature of pulsars?,” *Advances in Space Research*, vol. 40, 2007.
- [62] K. Wette *et al.*, “Searching for gravitational waves from cassiopeia a with ligo,” *Class. Quant. Grav.*, vol. 25, p. 235011, Dec. 2008.
- [63] J. Abadie *et al.*, “First search for gravitational waves from the youngest known neutron star,” *ApJ*, vol. 722, p. 1504, Oct. 2010.
- [64] P. R. Brady and T. Creighton, “Searching for periodic sources with LIGO. II: Hierarchical searches,” *Phys. Rev. D.*, vol. 61, p. 082001, 2000.
- [65] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, “Equation of State Calculations by Fast Computing Machines,” *J.Chem.Phys.*, vol. 21, pp. 1087–1092, June 1953.
- [66] S. Geman and D. Geman, “Stochastic relaxation, gibbs distributions, and the bayesian restoration of images,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. PAMI-6, pp. 721 –741, nov. 1984.
- [67] J. A. Nelder and R. Mead, “A simplex method for function minimization,” *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.
- [68] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, “Optimization by simulated annealing,” *Science*, vol. 220, no. 4598, pp. 671–680, 1983.
- [69] J. Skilling, “Nested sampling for general Bayesian computation,” *Bayesian Analysis*, vol. 1, no. 4, pp. 833–860, 2006.
- [70] C. Audet and J. E., “Mesh adaptive direct search algorithms for constrained optimization,” *SIAM Journal on optimization*, vol. 17, p. 2006, 2004.
- [71] M. A. Abramson, C. Audet, J. E. Dennis, and S. L. Digabel, “Orthomads: A deterministic mads instance with orthogonal directions,” *SIAM Journal on Optimization*, vol. 20, no. 2, pp. 948–966, 2009.
- [72] A. R. Conn and S. L. Digabel, “Use of quadratic models with mesh adpative direct search for constrained black box optimization,” *Optimization methods and software*, 2011.
- [73] M. Abramson, C. Audet, G. Couture, J. Dennis, Jr., S. Le Digabel, and C. Tribes, “The NOMAD project.” Software available at <http://www.gerad.ca/nomad>.
- [74] B. P. Abbott *et al.*, “Ligo: the laser interferometer gravitational-wave observatory,” *Reports on Progress in Physics*, vol. 72, p. 076901, July 2009.

- [75] T. Accadia *et al.*, “Status of the virgo project,” *Class. Quant. Grav.*, vol. 28, p. 114002, June 2011.
- [76] H. Grote, “The geo 600 status,” *Class. Quant. Grav.*, vol. 27, p. 084003, Apr. 2010.
- [77] J. Abadie *et al.*, “All-sky search for periodic gravitational waves in the full s5 ligo data,” *Phys. Rev. D.*, vol. 85, p. 022001, 2012.
- [78] B. Abbott *et al.*, “Searches for gravitational waves from known pulsars with science run 5 ligo data,” *ApJ*, vol. 713, p. 671, Apr. 2010.
- [79] B. Abbott *et al.*, “Einstein@home search for periodic gravitational waves in early s5 ligo data,” *Phys. Rev. D.*, vol. 80, p. 042003, Aug. 2009.
- [80] G. M. Harry, “Advanced ligo: the next generation of gravitational wave detectors,” *Class. Quant. Grav.*, vol. 27, p. 084006, Apr. 2010.
- [81] The Virgo Collaboration, “Advanced virgo baseline design,” 2009.
- [82] B. Willke *et al.*, “The geo-hf project,” *Class. Quant. Grav.*, vol. 23, p. 207, Apr. 2006.
- [83] M. Punturo *et al.*, “The third generation of gravitational wave observatories and their science reach,” *Class. Quant. Grav.*, vol. 27, p. 084007, Apr. 2010.
- [84] B. Knispel and B. Allen, “Blandford’s argument: The strongest continuous gravitational wave signal,” *Phys. Rev. D.*, vol. 78, p. 044031, Aug. 2008.
- [85] B. Abbott *et al.*, “Searches for periodic gravitational waves from unknown isolated sources and scorpius x-1: Results from the second ligo science run,” *Phys. Rev. D.*, vol. 76, p. 082001, 2007.
- [86] H. J. Pletsch, “A sliding coherence window technique for hierarchical detection of continuous gravitational waves,” *Phys. Rev. D.*, vol. 83, p. 122003, 2011.
- [87] S. Dhurandhar, B. Krishnan, H. Mukhopadhyay, and J. T. Whelan, “Cross-correlation search for periodic gravitational waves,” *Phys. Rev. D.*, vol. 77, p. 082001, Apr. 2008.
- [88] C. Cutler, I. Gholami, and B. Krishnan, “Improved stack-slide searches for gravitational-wave pulsars,” *Phys. Rev. D.*, vol. 72, no. 4, p. 042004, 2005.
- [89] R. Prix, “Search for Continuous Gravitational Waves:. Simple Criterion for Optimal Detector Networks,” in *The Eleventh Marcel Grossmann Meeting On Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories* (H. Kleinert, R. T. Jantzen, & R. Ruffini, ed.), p. 2441, 2008.

- [90] B. Knispel, *Pulsar Discoveries by Volunteer Distributed Computing and the Strongest Continuous Gravitational Wave Signal*. PhD thesis, Leibniz Universität Hannover, 2011.
- [91] P. Jaranowski, A. Królak, and B. F. Schutz, “Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection,” *Phys. Rev. D.*, vol. 58, p. 063001, 1998.
- [92] H. J. Pletsch and B. Allen, “Exploiting Large-Scale Correlations to Detect Continuous Gravitational Waves,” *Phys. Rev. Lett.*, vol. 103, p. 181102, Oct. 2009.
- [93] C. Messenger, “A semi-coherent search strategy for known continuous wave sources in binary systems,” *Phys. Rev. D.*, vol. 84, p. 083003, 2011.
- [94] K. Wette, “Estimating the sensitivity of wide-parameter-space searches for gravitational-wave pulsars,” *Phys. Rev. D.*, vol. 85, p. 042003, 2012.
- [95] G. Woan. (private communication).
- [96] P. R. Williams and B. F. Schutz, “An efficient matched filtering algorithm for the detection of continuous gravitational wave signals,” 1999.
- [97] P. Patel, X. Siemens, R. Dupuis, and J. Betzwieser, “Implementation of barycentric resampling for continuous wave searches in gravitational wave data,” *Phys. Rev. D.*, vol. 81, p. 084032, Apr. 2010.
- [98] R. Prix and M. Shaltev, “Search for continuous gravitational waves: Optimal stackslide method at fixed computing cost,” *Phys. Rev. D*, vol. 85, p. 084010, Apr 2012.
- [99] S. V. Dhurandhar and A. Vecchio, “Searching for continuous gravitational wave sources in binary systems,” *Phys. Rev. D.*, vol. 63, p. 122001, 2001.
- [100] K. Wette, “Estimating the sensitivity of wide-parameter-space searches for gravitational-wave pulsars,” *Phys. Rev. D*, vol. 85, p. 042003, Feb 2012.
- [101] S. Le Digabel, “Algorithm 909: Nomad: Nonlinear optimization with the mads algorithm,” *ACM Trans. Math. Softw.*, vol. 37, pp. 44:1–44:15, Feb. 2011.
- [102] “Einstein@home.” <http://www.einsteinathome.org>, Nov. 2011.
- [103] I. Harry, B. Allen, and B. Sathyaprakash, “A stochastic template placement algorithm for gravitational wave data analysis,” *Phys.Rev.D*, vol. 80, p. 104014, 2009.
- [104] M. Vallisneri, “Use and abuse of the fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects,” *Phys. Rev. D*, vol. 77, p. 042001, Feb 2008.

- [105] “lalsuite.” <https://www.lsc-group.phys.uwm.edu/daswg/projects/lalsuite.html>, Nov. 2011.
- [106] B. Abbott *et al.*, “All-sky search for periodic gravitational waves in LIGO S4 data,” *Phys. Rev.*, vol. D77, p. 022001, 2008.
- [107] A. Watts, B. Krishnan, L. Bildsten, and B. F. Schutz, “Detecting gravitational wave emission from the known accreting neutron stars,” *Mon. Not. Roy. Astron. Soc.*, vol. 389, p. 839, 2008.
- [108] S. Ghahramani, *Fundamentals of Probability*. Prentice Hall, Inc., 2000.
- [109] B. Abbott *et al.*, “All-sky search for periodic gravitational waves in ligo s4 data,” *Phys. Rev. D.*, vol. 77, p. 022001, 2008.
- [110] P. Jaranowski and A. Krolak, “Data analysis of gravitational-wave signals from spinning neutron stars. III: Detection statistics and computational requirements,” *Phys. Rev. D.*, vol. 61, p. 062001, 2000.
- [111] A. Krolak, M. Tinto, and M. Vallisneri, “Optimal filtering of the LISA data,” *Phys.Rev.*, vol. D70, p. 022003, 2004.
- [112] B. Abbott *et al.*, “Einstein@home search for periodic gravitational waves in ligo s4 data,” *Phys. Rev. D*, vol. 79, p. 022001, Jan 2009.
- [113] M. R. Chernick, “The jackknife: a resampling method with connections to the bootstrap,” *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 4, no. 2, pp. 224–226, 2012.
- [114] M. Shaltev and R. Prix, “Fully coherent follow-up of continuous gravitational-wave candidates,” *Phys. Rev. D*, vol. 87, p. 084057, Apr 2013.
- [115] P. Astone, M. Bassan, P. Bonifazi, K. Borkowski, R. Budzynski, *et al.*, “All-sky search of NAUTILUS data,” *Class.Quant.Grav.*, vol. 25, p. 184012, 2008.
- [116] I. Gholami, *Data analysis of continuous gravitational waves*. Nov. 2007.
- [117] B. Abbott *et al.*, “Setting upper limits on the strength of periodic gravitational waves from PSR J1939 + 2134 using the first science data from the GEO 600 and LIGO detectors,” *Phys. Rev. D.*, vol. 69, p. 082004, 2004.
- [118] D. T. Nguyen, M. Shaltev, and J. Ostermann, “Error concealment in the network abstraction layer for the scalability extension of h.264/avc,” *Proc. ICCE06, International Conference on Communications and Electronics*, pp. 274–278, Oct. 2006.
- [119] J. H. Conway and N. J. A. Sloane, *Sphere packings, lattices and groups*. A Series of Comprehensive Studies in Mathematics, Springer, 1999.

- [120] R. Prix and B. Krishnan, “Targeted search for continuous gravitational waves: Bayesian versus maximum-likelihood statistics,” Tech. Rep. LIGO-P0900066-v2, July 2009.
- [121] M. Shaltev, “Coherent follow-up of continuous gravitational-wave candidates: minimal required observation time,” *Journal of Physics: Conference Series*, vol. 363, no. 1, p. 012043, 2012.
- [122] J. Abadie *et al.*, “Calibration of the LIGO Gravitational Wave Detectors in the Fifth Science Run,” *Nucl.Instrum.Meth.*, vol. A624, pp. 223–240, 2010.
- [123] D. Tatsumi, R. Takahashi, K. Arai, N. Nakagawa, K. Agatsuma, *et al.*, “Current status of Japanese detectors,” *Class.Quant.Grav.*, vol. 24, pp. S399–S404, 2007.
- [124] D. MacLeod, S. Fairhurst, B. Hughey, A. Lundgren, L. Pekowsky, *et al.*, “Reducing the effect of seismic noise in LIGO searches by targeted veto generation,” *Class.Quant.Grav.*, vol. 29, p. 055006, 2012.
- [125] M. Nayyar and B. J. Owen, “R-modes of accreting hyperon stars as persistent sources of gravitational waves,” *Phys.Rev.*, vol. D73, p. 084001, 2006.
- [126] R.-X. Xu, “To probe into pulsar’s interior through gravitational waves,” *Astropart.Phys.*, vol. 25, pp. 212–219, 2006.
- [127] L. Stella, S. Dall’Osso, G. Israel, and A. Vecchio, “Gravitational radiation from newborn magnetars,” *Astrophys.J.*, vol. 634, pp. L165–L168, 2005.
- [128] S. Johnston, G. Hobbs, S. Vigeland, M. Kramer, J. Weisberg, *et al.*, “Evidence for alignment of the rotation and velocity vectors in pulsars,” *Mon.Not.Roy.Astron.Soc.*, vol. 364, pp. 1397–1412, 2005.
- [129] O. Benhar, “Neutron star matter equation of state and gravitational wave emission,” *Mod.Phys.Lett.*, vol. A20, pp. 2335–2350, 2005.
- [130] A. Passamonti, E. Gaertig, and K. Kokkotas, “Gravitational waves from the evolution of the f-mode instability in neutron stars,” 2012.
- [131] G. Rupak and P. Jaikumar, “The r-mode instability in strange stars with a crystalline crust,” 2012.
- [132] P. D. Lasky, B. Zink, and K. D. Kokkotas, “Gravitational Waves and Hydromagnetic Instabilities in Rotating Magnetized Neutron Stars,” 2012.
- [133] I. Santiago-Prieto, I. Heng, D. Jones, and J. Clark, “Prospects for transient gravitational waves at r-mode frequencies associated with pulsar glitches,” *J.Phys.Conf.Ser.*, vol. 363, p. 012042, 2012.

- [134] S. S. Yazadjiev and D. D. Doneva, “Possible dark energy imprints in gravitational wave spectrum of mixed neutron-dark-energy stars,” *JCAP*, vol. 1203, p. 037, 2012.
- [135] W. Kastaun, “Nonlinear Decay of r modes in Rapidly Rotating Neutron Stars,” *Phys.Rev.*, vol. D84, p. 124036, 2011.
- [136] E. Gaertig, K. Glampedakis, K. D. Kokkotas, and B. Zink, “The f-mode instability in relativistic neutron stars,” *Phys.Rev.Lett.*, vol. 107, p. 101102, 2011.
- [137] G. Burgio, V. Ferrari, L. Gualtieri, and H.-J. Schulze, “Oscillations of hot, young neutron stars: Gravitational wave frequencies and damping times,” *Phys.Rev.*, vol. D84, p. 044017, 2011.
- [138] A. Corsi and B. J. Owen, “Maximum gravitational-wave energy emissible in magnetar flares,” *Phys.Rev.*, vol. D83, p. 104014, 2011.
- [139] M. Gabler, U. Sperhake, and N. Andersson, “Non-linear radial oscillations of neutron stars,” *Phys.Rev.*, vol. D80, p. 064012, 2009.
- [140] C. Cutler, “An Improved, ‘phase-relaxed’ F-statistic for gravitational-wave data analysis,” *Phys.Rev.*, vol. D86, p. 063012, 2012.
- [141] B. Haskell, D. Jones, and N. Andersson, “Mountains on Neutron Stars: Accreted vs. Non-Accreted crusts,” *Mon.Not.Roy.Astron.Soc.*, vol. 373, pp. 1423–1439, 2006.
- [142] C. Cutler, “Gravitational waves from neutron stars with large toroidal B fields,” *Phys.Rev.*, vol. D66, p. 084025, 2002.
- [143] C. Cutler and K. S. Thorne, “An Overview of gravitational wave sources,” 2002.
- [144] C. Cutler and D. I. Jones, “Gravitational wave damping of neutron star wobble,” *Phys.Rev.*, vol. D63, p. 024002, 2001.
- [145] G. Ushomirsky, L. Bildsten, and C. Cutler, “Gravitational waves from low mass X-ray binaries: A Status report,” 1999.
- [146] D. R. Lorimer, “Binary and millisecond pulsars,” *Living Reviews in Relativity*, vol. 8, no. 7, 2005.
- [147] G. Hobbs, W. Coles, R. Manchester, and D. Chen, “Developing a pulsar-based timescale,” 2010.
- [148] V. M. Kaspi, “Grand Unification in Neutron Stars,” *Proc.Nat.Acad.Sci.*, vol. 107, pp. 7147–7152, 2010.

- [149] M. Zimmermann and E. Szedenits, “Gravitational waves from rotating and precessing rigid bodies: Simple models and applications to pulsars,” *Phys. Rev. D*, vol. 20, pp. 351–355, Jul 1979.
- [150] S.-N. Zhang and Y. Xie, “Why do the braking indices of pulsars span over more than 100 millions?,” 2012.
- [151] A. Biryukov, G. Beskin, and S. Karpov, “Monotonic and cyclic components of radio pulsars spin-down,” *Mon.Not.Roy.Astron.Soc.*, vol. 420, pp. 103–117, 2012.
- [152] J. P. Ridley and D. R. Lorimer, “Isolated pulsar spin evolution on the P-Pdot Diagram,” 2010.
- [153] M. A. Livingstone, V. M. Kaspi, and E. V. Gotthelf, “A braking index for the young pulsar at the center of the supernova remnant kes 75,” *Astrophys.J.*, 2006.
- [154] M. Allen and J. Horvath, “Implications of a constant observed braking index for young pulsar’s spindown,” 1997.
- [155] H. J. Pletsch, “Sliding coherence window technique for hierarchical detection of continuous gravitational waves,” *Phys.Rev.*, vol. D83, p. 122003, 2011.
- [156] R. Prix and K. Wette, “Estimating sensitivity of the Einstein@Home search S5R5,” *LIGO DCC T1200272*, 2012.
- [157] A. Sesana, A. Vecchio, and C. N. Colacino, “The stochastic gravitational-wave background from massive black hole binary systems: implications for observations with Pulsar Timing Arrays,” 2008.
- [158] D. Lorimer, *Handbook of Pulsar Astronomy*. Cambridge Observing Handbooks for Research Astronomers, Cambridge University Press, 2004.
- [159] C. Palomba, “Pulsars ellipticity revised,” Dec. 1999.
- [160] L. Wade, X. Siemens, D. Kaplan, B. Knispel, and B. Allen, “Continuous Gravitational Waves from Isolated Galactic Neutron Stars in the Advanced Detector Era,” 2012.

Acknowledgment

I am grateful to Bruce Allen and Andrzej Krolak for refereeing my thesis.

I am grateful to Bruce Allen for the questions, advice and interest in my work and the given chance to be part of the Observational Relativity and Cosmology group at AEI and the International Max Planck Research School on Gravitational Wave Astronomy (IMPRS). His support in the last 4 years allowed me to contribute to the gravitational-wave community, to be part of the LIGO-Virgo collaboration and meet some of the most outstanding scientists in this research area.

I would like to thank the chairs of the Pulgroup Keith Riles, Graham Woan, Andrzej Krolak, Cristiano Palomba and Pia Astone for their leadership and useful comments and questions during the numerous telecons and face-to-face collaboration meetings.

I am grateful to Maria Alessandra Papa (MAP) and Badri Krishnan for their questions and suggestions. The weakly Continuous Waves sub-group meeting at AEI chaired by MAP has always been very stimulating and helping to move forward with my research. I am thankful for numerous discussions and advice of my AEI colleagues, i.a., Paola Leaci, Holger Pletsch, Karl Wette, Gian Mario Manca, Evan Götz, Thomas Dent, Francesco Salemi, Hyunjoo Kim, Christian Röver, Chris Messenger, Ben Sheard. I also want to thank to Oliver Bock and Bernd Machenschalk for some computing related help.

The inhabitants of office 033, Irene Di Palma, Giulio Mazzolo and the far away associate Letizia Sammut, I would like to thank for the time spent together. The first two are responsible for my Italian connection. It is *causa perduta* to acknowledge all of the Italian colleges, just to list a few Marco Drago, Salvatore Vitale and Riccardo Sturani. I would like to thank Vladimir Dergachev for all of the extraordinary interesting conversations and suggestions. I appreciated the discussions about science and life with the fellow PhD students Pablo Rosado, Tito Dal Canton, Berit Behnke, David Keitel and a person named Shaon Ghosh.

I am grateful to Alben Draycheva for a considerable amount of enlightening.

I am thankful to Carsten Aulbert and Henning Fehrmann for the constant support concerning not only the usage of the ATLAS cluster, but also my doctoral studies in general. In fact its due to Henning that I entered the AEI, helping with the cluster administration 6 years ago. I truly enjoyed that time.

I am grateful to Marian Ivanov and Dilyana Ivanova for their support and interest in my work. I am grateful to my parents Rositza and Valentin for their support and encouragement to do and study what I want.

My scientific progress in the last 4 years have been primarily driven by my supervisor Reinhard Prix. I am deeply grateful for the given freedom in research, the uncountable scientific discussions and advice. Thank you very much Reinhard!

Curriculum Vitae

Miroslav Shaltev

Name

2009 - 2013

Doctoral Studies, International Max Planck Research School on Gravitational Wave Astronomy, Max Planck Institute for Gravitational Physics (Albert-Einstein-Institute Hannover) and Leibniz Universität Hannover, Germany

2008 - 2009

Diploma in Physics, Lazer Zentrum Hannover, Leibniz Universität Hannover, Germany

2005 - 2008

Diploma studies in Physics, Leibniz Universität Hannover, Germany

2002 - 2005

Diploma studies in Technical Physics, Leibniz Universität Hannover, Germany

2001 - 2002

Diploma studies in Industrial Engineering, Leibniz Universität Hannover, Germany

1999 - 2000

Diploma studies in Engineering Physics, Sofia University “St. Kliment Ohridski”, Bulgaria

1999

Secondary High School “Ivan Vazov”, Dobrich, Bulgaria

May 01, 1981

Date of birth

Sofia

Place of birth