Upper Limits for Galactic Transient Sources of Gravitational Radiation From LIGO First Observations

by

Julien Sylvestre

Submitted to the Department of Physics in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

[September 2002]
June 2002

© Julien Sylvestre, MMII. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author

Julien Sylvestre

Department of Physics
June 24, 2002

Certified by

Rainer Weiss
Professor
Thesis Supervisor

Accepted by

Thomas J. Greytak
Associate Department Head for Education
Upper Limits for Galactic Transient Sources of Gravitational
Radiation From LIGO First Observations

by

Julien Sylvestre

Submitted to the Department of Physics
on June 24, 2002, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

A complete, scalable prototype analysis system for the observation of gravitational
wave bursts with a network of interferometric detectors is designed and tested. This
system detects localized transients in the gravitational wave data streams from two
independent detectors using a time-frequency algorithm, applies veto conditions on
these transients based on auxiliary channels, and looks for significant coincidences
between the surviving transients from the two detectors. The analysis system was
optimized for setting rate upper limits on three populations of astronomical sources,
using preliminary data from the LIGO Project (the E7 Engineering Run data). The
three classes of sources were core collapses, bar-mode instabilities in neutron stars,
and equal-mass black hole binary coalescences, and a 95% confidence rate upper
limit of 2 per hour was set, for sources uniformly distributed within a distance of 6
pc, 950 pc, and 3.2 kpc from the Earth, respectively. A detailed discussion of the
character of the noise in the data used to derive these limits shows that a network of
interferometers including the instruments of the GEO, LIGO and VIRGO Projects
should attain a 50% detection efficiency for these sources out to distances of 40 kpc,
1 Mpc and 40 Mpc, respectively, for a false alarm rate giving expected upper limits
of 6, 0.6, and $2 \cdot 10^{-4}$ per year per galaxy, respectively, for a full year of observation.

Thesis Supervisor: Rainer Weiss
Title: Professor
Acknowledgments

It is a pleasure to thank everybody who has helped me in getting to the point where I could write this thesis.

I had the chance to be a member of the Gravitation and Cosmology Research Group at MIT, and to exchange on a regular basis with many of the outstanding experimentalists who built the LIGO instruments. Among the members of this group, I am especially grateful to my thesis supervisor for his inspiring, rigorous and liberal guidance.

I also enjoyed the much needed encouragement and care from my family throughout my residence at MIT. My beloved wife, by her unconditional and comforting support, did much to help bring this work to its reality. It is with a deep gratitude that I dedicate this thesis to her.
# Contents

Introduction 13

1 Gravitational Waves and Their Detection 17
  1.1 Propagation .......................... 17
  1.2 Generation .......................... 21
  1.3 Detection ............................ 24
  1.4 Interferometric Detectors ........... 25

2 Transient Sources of Gravitational Radiation 33
  2.1 Black Hole Binary Coalescences ........ 33
  2.2 Core Collapses ....................... 39
  2.3 Neutron Star Instabilities ........... 43

3 Analysis Pipeline 49
  3.1 Data Selection ........................ 51
  3.2 The TFCLUSTERS Analysis ............ 52
    3.2.1 TFCLUSTERS: Algorithm .......... 52
    3.2.2 TFCLUSTERS: Implementation ...... 54
  3.3 Veto Generation ...................... 58
    3.3.1 PSLGlitch ........................ 59
    3.3.2 MICHGlitch ....................... 61
  3.4 Pipeline Implementation ............. 63
# List of Figures

1-1 The effect of a gravitational wave. ........................................ 20
1-2 The power recycled Fabry-Perot interferometer. ....................... 29
1-3 Designed sensitivity of the LIGO 4 km interferometers. ............... 30

2-1 Theoretical distribution of the mass of compact objects in binaries. . 38
2-2 The strain amplitude for three Zwerger-Müller waveforms. ........... 42

3-1 Flow diagram of the analysis pipeline. ..................................... 50
3-2 Black pixel probability for the estimated power thresholds. .......... 57
3-3 PSL signal during a glitch. .................................................. 60
3-4 Eigenvalues of the component analysis for the PSL glitch. ............. 61
3-5 First component for the PSL glitch. ....................................... 62

4-1 Example of an acceptance region. ......................................... 72
4-2 Test of an acceptance region with simulated data. ...................... 73

5-1 Power spectral density for LLO-4k. ....................................... 81
5-2 Power spectral density for LHO-2k. ....................................... 82
5-3 The ratio of the total observation time to the livetime. ................. 85
5-4 The number of coincident events in the Playground dataset (A). .... 85
5-5 The number of coincident events in the Playground dataset (B). .... 86
5-6 The number of coincident events in the Playground dataset (C). .... 86
5-7 95% confidence level projected upper limit (A). ....................... 87
5-8 95% confidence level projected upper limit (B). ....................... 87
5-9 95% confidence level projected upper limit (C). ....................... 88
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5-10</td>
<td>The effect of the vetoes on events power</td>
<td>89</td>
</tr>
<tr>
<td>5-11</td>
<td>The probability of detection (POD) for various signals</td>
<td>91</td>
</tr>
<tr>
<td>5-12</td>
<td>Probability of detection vs. signal coherence</td>
<td>93</td>
</tr>
<tr>
<td>5-13</td>
<td>The 95% confidence region</td>
<td>95</td>
</tr>
<tr>
<td>5-14</td>
<td>Probability of detection for the bar-mode instability</td>
<td>96</td>
</tr>
<tr>
<td>5-15</td>
<td>Efficiency for the bar-mode instability</td>
<td>97</td>
</tr>
<tr>
<td>5-16</td>
<td>Upper limit for the bar-mode instability</td>
<td>97</td>
</tr>
<tr>
<td>5-17</td>
<td>Probability of detection for core-collapses</td>
<td>99</td>
</tr>
<tr>
<td>5-18</td>
<td>Upper limit on the rate of core collapses</td>
<td>99</td>
</tr>
<tr>
<td>5-19</td>
<td>Probability of detection for black hole mergers</td>
<td>101</td>
</tr>
<tr>
<td>5-20</td>
<td>Upper limit for black hole mergers</td>
<td>101</td>
</tr>
<tr>
<td>5-21</td>
<td>Time delay histogram</td>
<td>103</td>
</tr>
<tr>
<td>5-22</td>
<td>Distribution of number of events per segment (LLO-4k)</td>
<td>103</td>
</tr>
<tr>
<td>5-23</td>
<td>Distribution of number of events per segment (LHO-2k)</td>
<td>104</td>
</tr>
<tr>
<td>5-24</td>
<td>Simulated distribution of number of coincidences</td>
<td>107</td>
</tr>
<tr>
<td>5-25</td>
<td>Mean to variance ratio vs. variance of number of events</td>
<td>109</td>
</tr>
</tbody>
</table>
List of Tables

5.1 Fraction of segments satisfying the data cuts. .......................... 80
5.2 Configurations used to optimize TFCLUSTERS. ......................... 84
5.3 Number of events in the analysis pipeline. .............................. 94
5.4 Number of coincidences per segment. ................................. 105
5.5 Comparison of the distributions of event parameters. ................ 110
5.6 Projected upper limits. ................................................. 113
5.7 Projected, rescaled upper limits. .................................... 115
Introduction

A fascinating prediction of the theory of General Relativity is that dynamical perturbations of space-time, gravitational waves, can propagate at the speed of light over cosmological distances, with practically no attenuation. These waves couple only weakly with matter, and hence require extreme conditions to be generated with sufficient amplitude to be detectable; such conditions are only realized in highly relativistic astronomical systems. While beautiful indirect proofs of the existence of gravitational waves exist (e.g. from the observation of the Hulse-Taylor pulsar [1]), a direct detection of a wave, through the observation of its stretching of a detector, is still missing, despite four decades of experimentation with resonant bar detectors. However, the radical improvements in sensitivity and bandwidth brought by a new generation of detectors, large laser interferometers, might be sufficient to open the gravitational wave window on the Universe. Four large collaborations [2] are currently commissioning six interferometers in North America, Europe and Asia, with plans for significant upgrades [3, 4] and new instruments [6, 5] over the next decade.

A few direct detections would bring a wealth of new information about the theory of General Relativity in the strong field regime, about the dynamics and the astrophysics of the systems that radiated the waves, and possibly about other fundamental aspects of the Universe (such as the mass of the neutrinos [7] or the value of the Hubble constant [8]). While certainly less interesting, the absence of any detection could alternatively be used to set upper limits on the rate or amplitude of gravitational waves bathing the Earth, thus constraining models of the formation of sources of gravitational radiation, or models of their dynamics. The goal of this thesis will be to design, implement and test the required tools and supporting concepts that are
necessary to set such limits. The tests will be performed using preliminary data from two of the interferometers of the LIGO Project; consequently, the upper limits that will be measured are not expected to be interesting in themselves. Rather, the focus of this work will be on carefully designing and optimizing the analysis pipeline, keeping in mind that the sensitivity of the instruments should improve sufficiently over a time-scale of less than a year to give a few orders of magnitude improvement in the strain sensitivity of the interferometers.

The data used in this work were collected between 2001 December 28 18:00UTC and 2002 January 14 18:00UTC by the LIGO Project, a data collection period known as the E7 Engineering Run. Data from two interferometers were used: the 2 km-arm interferometer in Hanford, Washington (the LHO-2k interferometer), and the 4 km-arm interferometer in Livingston, Louisiana (the LLO-4k interferometer). Details about the instruments and their interaction with a gravitational wave are given in chapter 1.

The E7 data were searched for a number of specific types of signals, corresponding to a subset of the many sources of gravitational radiation that have been analyzed as candidates for detection by interferometric detectors (cf. [9] for a recent review). These sources can be classified into two distinct populations, based on the typical duration of their gravitational wave signal. Sources which emit steady signals with amplitudes varying over time-scales of months or years form the first population. Examples of such sources include rotating, deformed neutron stars [10] which emit a quasi-periodic signal, and the density fluctuations of the early universe that may have produced a stochastic background of gravitational radiation that is still detectable [11]. Only sources from the second population of shorter, transient gravitational wave signals will be considered in any detail here. The principal reason behind this restriction of scope lies in the vast differences in the requirements for the analysis methods that are used to detect long and short signals.

Three specific populations of transient sources will be studied, and upper limits on the rate of occurrence of transients from each population will be set. As described in detail in chapter 2, these are coalescing binaries consisting of two equal mass black
holes, collapsing iron cores in type II supernovae, and unstable bar-modes in newborn neutron stars. These sources were chosen because they offer a good variety of waveforms, and because they enjoy a strong theoretical bias among transient sources as being the most likely candidates for LIGO first detections. A noteworthy exception with respect to the latter criterion is the collapse of neutron star binaries; this particular source was not considered here because special optimal techniques (matched filtering) can be developed to search for its signals in the data, and these techniques are different from the ones that are used in this work.

The principal characteristic of the search algorithm used here that differentiates it from other approaches is its inherent robustness against poor prior knowledge of the signal. Since the details of the waveforms from the three sources mentioned above are not known, it is important both for the detection and for the computation of upper limits to have a search algorithm that is relatively independent of those details, in the sense that the probability of detection for a certain signal is only a weak function of perturbations of that signal. The TFCLUSTERS search algorithm, which is based on clustering analysis in the time-frequency plane, possesses this required robustness, and will be used in this work. It is described in chapter 3, together with the supporting infrastructure of the analysis pipeline. The most important component of this infrastructure consists in the algorithms that are used to generate veto conditions to reject spurious bursts in the interferometers, based on information from parts of the system that are not sensitive to gravitational waves. Two vetoes were defined, one for each of the two interferometers, and they both proved to be useful at improving the quality of the measured upper limits.

The output of the analysis pipeline is the number of coincident events observed during a certain time. A careful description of the operations required to transform this number into a rate upper limit for a given class of sources is presented in chapter 4. Essentially, this process involves the estimates of the number of background events, of the livetime, and of the efficiency of the pipeline at detecting sources picked at random from this class. The efficiency is estimated by performing numerical simulations where signal models from chapter 2 are injected into the E7 data. The background, on the
other hand, is estimated by adding a time lag to the LLO-4k data stream, and by using the analysis pipeline to compute the number of resulting coincidences, which cannot be the result of the presence of astrophysical signals in the data.

The analysis pipeline, as can be expected, has a number of parameters that can be adjusted to the characteristics of the data to provide the best detection possibilities, or the smallest upper limits. The parameters of the pipeline will be optimized for the E7 data, with a method that is general enough to be reusable on other datasets. A discussion of this optimization is presented in the first sections of chapter 5.

The upper limits on the rate of occurrence within a certain volume of each of the three classes of objects mentioned above are also presented in chapter 5. Not surprisingly given the quality of the E7 data, these upper limits, while novel in the way they were obtained, do not bring new astrophysical information forward. However, they were used to generate a number of criteria for data quality which, if all met in the final stage of commissioning of the instruments, would give limits that are challenging some theoretical models. These limits are discussed in regards of the robustness of the analysis pipeline by analyzing the outcomes of numerical simulations with randomly perturbed waveforms.

Finally, the validity of the whole analysis is also discussed in chapter 5. The most worrisome point appears to be related to the lack of stationarity of the data on time-scales of the order of tens of minutes. Although there is no direct support of this from the data at hand, numerical simulations using a model of the non-stationarity of the actual data as input show that it is possible that the coincident events may not be accurately modeled by a Poisson process. It is shown, however, that a modest improvement of the data stationarity in future data runs should guarantee the required convergence to a Poisson process.
Chapter 1

Gravitational Waves and Their Detection

"Peut-être l'immobilité des choses autour de nous leur est-elle imposée par notre certitude que ce sont elles et non pas d’autres, par l'immobilité de notre pensée en face d'elles.”

Marcel Proust

Analogously to accelerated charged particles that emit electro-magnetic waves, accelerated massive particles emit gravitational waves. These waves are deformations of spacetime that propagate at the speed of light. Their effect on an observer is to modify his local measurements of distance between freely-falling bodies. The detection of these distance changes is a very challenging technological problem, even for the strongest waves produced in astronomical systems with extreme densities and internal velocities.

1.1 Propagation

The equations of General Relativity in vacuum, when linearized, admit a solution in the form of a plane wave propagating with the speed of light (see chapter 35 of [12]
for details). The linearized metric is written as

\[ g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} + O(h_{\mu\nu}^2), \]  

(1.1)

where \( \eta_{\mu\nu} \) is the Minkowski metric, and the symmetric tensor \( h_{\mu\nu} \) is the small metric perturbation that will be interpreted below as gravitational waves.

Written in this manner, \( h_{\mu\nu} \) appears to have ten degrees of freedom; however, eight of these can be removed by applying appropriate gauge conditions. The *transverse-traceless* gauge requires that (i) only the spatial components of \( h_{\mu\nu} \) are non-zero, (ii) the spatial components have a vanishing divergence, and (iii) the spatial components have a vanishing trace\(^1\). The first two conditions make the wave transverse to its direction of propagation, hence the name for this particular gauge. In this gauge, the vacuum solution to Einstein field equations for Eq. (1.1) for a wave propagating in the \( z \) direction is

\[ h_{xx}(t) = -h_{yy}(t) = h_+ \cos[\omega(ct - z)] \] 

(1.2)

\[ h_{xy}(t) = h_{yx}(t) = h_\times \cos[\omega(ct - z)], \] 

(1.3)

where the two polarizations of the wave are explicitly shown, and \( c \) is the speed of light.

For two freely-falling particles \( A \) and \( B \), the spatial coordinates \( r^B_j \) of particle \( B \), in a Lorentz frame with particle \( A \) at its origin, change under the acceleration from an incident gravitational wave as:

\[ \frac{d^2 r^B_j}{dt^2} = \frac{1}{2} \sum_k \left( \frac{\partial^2 h_{jk}}{\partial t^2} \right) r^B_k. \] 

(1.4)

Accordingly, if the two test particles are initially at rest with respect to each other,

---

\( ^1 \)The conditions (i)-(iii) eliminate six degrees of freedom. Of the four remaining ones, two are eliminated in the plane-wave solution (normalization of the wave vector and orthogonality of the amplitude and wave vectors).
their separation varies according to:

\[ r_j^B(t) = r_j^B(t_0) \sum_k \left[ \delta_{jk} + \frac{h_{jk}(t)}{2} \right]. \tag{1.5} \]

Because of this, \( h_{jk} \) is called the strain amplitude of the gravitational wave. Figure 1-1 illustrates the effect of the + polarization of a gravitational wave on a ring of test particles; the \( \times \) polarization has a similar effect, excepted for a rotation of the flow pattern by 45°.

It can be shown [12] that an arbitrary wave can always satisfy the conditions of the transverse-traceless gauge\(^2\). Hence, the general form of a gravitational wave in otherwise flat spacetime can be written as:

\[ h(t) = h_+(t)e_+ + h_\times(t)e_\times, \tag{1.6} \]

where the two polarization tensors are defined by

\[ e_+ = e_x \otimes e_x - e_y \otimes e_y \tag{1.7} \]
\[ e_\times = e_x \otimes e_y + e_y \otimes e_x, \tag{1.8} \]

for orthonormal basis vectors \( e_x \) and \( e_y \), and a wave propagating in the \( z \) direction.

This work will be concerned primarily with the detection of the waves in the weak field at Earth, and therefore only the physics relevant to their propagation in flat spacetime will be discussed. An interesting review of the problems related to their propagation in strongly curved spacetimes and the transition to flat spacetime can be found in [13].

\(^2\)This is only true for waves that are solution of the Einstein field equation in vacuum; i.e., it is not true for waves in a strongly curved spacetime.
Figure 1-1: The red dots show the position of non-interacting particles initially forming a ring at different phases of a + polarization, monochromatic gravitational wave normally incident on the ring. The phase angle is labeled in the center of the rings. The arrows represent the magnitude of the velocity of the particles.
1.2 Generation

Gravitational waves are inherently weak, and consequently they can only be emitted in amounts large enough for detection if the particles of the source are accelerated coherently, so that their individual contributions add up to build a significant wave. This naturally suggests the study of the time-evolution of the low-order moments of the mass distribution of the source.

From dimensional analysis, it is simple to show that if the dimensionless strain amplitude of gravitational waves from a given source is only related to the gravitational constant $G$, $c$, and some power of the $n^{\text{th}}$ time derivative of the $m^{\text{th}}$ mass moment of the source, then $n = m$. The first (second) derivative of the first (second) mass moment is zero by conservation of total mass (total linear momentum). Hence, the lowest order moment for the radiation of gravitational waves is the third moment, also known as the mass quadrupole moment. A similar dimensional analysis for the $n^{\text{th}}$ derivative of the $m^{\text{th}}$ current moment gives $(m, n) = (1, 1)$ and $(2, 2)$ as the two lowest order candidates. The first derivative of the first current moment is zero if no external forces act on the source (or, again, if total linear momentum is conserved), leaving the second derivative of the second current moment as the lowest order current moment for the radiation of gravitational waves. This moment is also called the mass quadrupole current moment.

The formal analysis, when carried out in the nearly Newtonian approximation\(^3\), leads to the following formula \([14]\), which agrees with the results of the above dimensional analysis:

$$h_{jk}^{TT}(t, x) = \left[ \frac{2}{c^6} \frac{d^2}{dt^2} \mathcal{I}_{jk}(ct - r) + \frac{8}{3d^3} \epsilon_{pql} \frac{d^2}{dt^2} S_{kp}(ct - r)n_q \right]^{TT} \quad (1.9)$$

Here, $\mathcal{I}_{jk}(ct - r)$ and $S_{kp}(ct - r)$ are the mass and current quadrupole moments evaluated at the appropriate retarded time ($r = |x|$), $d$ is the distance to the center-of-mass of the source, $n$ is the unit vector in the direction of the source, $\epsilon$ is the antisymmetric

\(^{3}\text{i.e. for sources with slow internal velocities} \ll c \text{ and weak gravitational potentials} \ll GM/R.\)
tensor, and \( TT \) denotes the transverse-traceless projection. Most sources’ emission (with the notable exception of the r-mode instability in neutron stars) will be dominated by the mass quadrupole term; hence the preponderance of this term in the discussion below.

Although gravitational waves couple only weakly to matter, they can have a significant effect on the dynamics of the source. A classical example of this is PSR1913+16, a binary pulsar system with a slowly decaying orbit due to the loss of energy to gravitational wave emission [1]. The energy flux for a wave is obtained from the stress-energy tensor,

\[
T_{tr} = \frac{1}{32\pi} \frac{c^4}{G} \langle h_{jkr}^{TT} h_{jkr}^{TT} \rangle,
\]

where \( \langle \cdot \rangle \) means “averaged over many cycles”. Substituting the mass quadrupole part of Eq.(1.9) into Eq.(1.10) and integrating over all directions gives the source luminosity in gravitational radiation:

\[
L = \frac{1}{5} c^5 \langle \mathcal{E}_{jk}^{(3)} \mathcal{E}_{jk}^{(3)} \rangle,
\]

where the parenthesized upper indices denote time derivatives (third, in this case). Similarly, the rate of change of the source angular momentum due to gravitational radiation is:

\[
\frac{dJ_i}{dt} = -\frac{2}{5} c^5 \varepsilon_{ijk} \langle \mathcal{M}_{jm}^{(3)} \mathcal{M}_{km}^{(3)} \rangle.
\]

The quadrupole approximation as described above is especially convenient for rough estimates of the magnitude of the effects involving gravitational waves. First,

\[
h_{jk}^{TT} \sim \frac{2}{d} c^4 M_{\text{NS}} \left( \frac{R}{T} \right)^2,
\]

where the second time derivative of the mass quadrupole moment is approximated by the kinetic energy in the non-spherical part of the source, which alone contributes to the emission of gravitational radiation (\( M_{\text{NS}} \) is the mass in that portion of the source, \( R \) and \( T \) are the characteristic length and time scales). Obviously, sources with a large deformation from sphericity will radiate the most; for those sources with
$M_{\text{NS}} \rightarrow M$, for $M$ the total source mass, Eq. (1.13) can be rewritten as

$$h_{ij}^{TT} \sim \frac{r_{\text{Sch}}}{d} \left( \frac{v}{c} \right)^2,$$

(1.14)

for $r_{\text{Sch}}$ the Schwarzschild radius of the source, and $v$ its characteristic internal velocity. This illustrates the second condition, after being massively non-spherical, for a strong radiator: the source must be relativistic (this, of course, is still true of sources with $v \sim c$ where the quadrupole approximation fails). It also illustrates the immense technological challenge presented by the detection of gravitational waves: even for $v \sim c$,

$$h \lesssim 10^{-19} \left( \frac{M}{1 \, M_{\odot}} \right) \left( \frac{1 \, \text{Mpc}}{d} \right).$$

(1.15)

It appears from Eq. (1.14) that massive objects give stronger wave amplitudes. The frequency of the waves, however, scales inversely with the mass of the source, thus limiting the observable mass range for a detector that is limited in bandwidth. Since the motions of particles in the source have to be coherent in order to get a strong radiator, waves with a period smaller than the light travel time around half the source circumference will be strongly suppressed. This leads to an upper bound in the wave frequency for a source of mass $M$ [15],

$$f \lesssim \frac{1}{2\pi} \frac{c}{\pi r_{\text{Sch}}} \sim 30 \, \text{kHz} \left( \frac{1 \, M_{\odot}}{M} \right).$$

(1.16)

If an object radiates monochromatic waves at frequency $f$, as in Eq. (1.3), then Eq. (1.10) gives the flux [13]

$$T_r = \frac{\pi}{4} \frac{c^3}{G} f^2 \langle h^2 \rangle \sim 30 \, \text{W/m}^2 \left( \frac{f}{1 \, \text{kHz}} \right)^2 \left( \frac{\langle h^2 \rangle}{10^{-40}} \right),$$

(1.17)

where $\langle h^2 \rangle = \langle h_+^2 + h_{x}^2 \rangle$, and the frequency and mean square amplitude are chosen to correspond to a realistically detectable source for LIGO. It should be noted that this energy flux is enormous, but that the source generally only radiates for a brief period of time at such a high flux. For instance, a source at distance $d$ that radiates
waves in all directions with mean square amplitude \( \langle h^2 \rangle \) at frequency \( f \) can radiate a significant fraction of its mass-energy in a short time, its luminosity being:

$$L = 2 \frac{M_\odot c^2}{s} \left( \frac{f}{1\text{ kHz}} \right)^2 \left( \frac{\langle h^2 \rangle}{10^{-40}} \right) \left( \frac{d}{1\text{ Mpc}} \right)^2.$$  \hspace{1cm} (1.18)

Since such a source with \( f = 1\text{ kHz} \) cannot be more massive than \( 30M_\odot \) [cf. Eq.(1.16)], and since the energy available to generate gravitational waves is certainly less than the kinetic energy of rotation of the source, which is assumed to spin at frequency \( f/2 \) (quadrupolar gravitational waves are emitted at twice the rotation frequency), the time duration for which the source can radiate is bounded by \( L\Delta t \lesssim \pi^2 M r_{\text{Sch}}^2 f^2 \), i.e. this source cannot radiate for more than \( \sim 1 \) second.

### 1.3 Detection

As it was discussed in section 1.1, an arbitrary gravitational wave propagating in flat spacetime can always be expressed in the transverse-traceless gauge; in this gauge, only the spatial components of the wave are non-zero. Consequently, all practical detectors are based on the precise measurement of changes in length due to the passage of gravitational waves.

Historically, the first attempts to detect gravitational waves were made with resonant bar detectors; modern versions of these apparatus consist in a massive solid bar that is isolated mechanically from its environment, and is cooled to cryogenic temperatures. A wave incident on the bar drives its normal modes; the resulting mechanical oscillations are then measured directly with an appropriate transducer. The best response is expected for the first longitudinal normal mode of the bar (which is symmetric with respect to a node at the center of the bar), which is generally arranged to have an eigenfrequency in the kilohertz range to match the astrophysical bias about source characteristic frequencies. The cross-section of these bars to gravitational waves is small, \( \sim 10^{-23} \text{ m}^2 \) \[12, \text{ Eq. (37.22)}\], and therefore a source like the one used in the example of Eq.(1.17) could only deposit energy in the bar at a rate of
\( \sim 10^{-22} \text{ W}, \) assuming the bar resonance frequency is within a few Hertz of 1 kHz. For a wave duration \( \Delta t \), and assuming that the bar dissipates energy with a characteristic time much longer than \( \Delta t \), this source could be detected if \( kT \lesssim \Delta t \cdot 10^{-22} \text{ W} \), for \( k \) the Boltzmann's constant, and \( T \) the temperature of the bar. This justifies the use of cryogenic temperatures, as \( T \lesssim 10 \text{ K}(\Delta t/1\text{s}) \).

There is now a well-developed worldwide network of cryogenic resonant bar detectors operated in coincidence[16], with resonance frequencies ranging from 694 Hz to 930 Hz. Up to five detectors on three continents were used in coincidence for 15.5 days (four or more instruments) during 1997 and 1998 to set with 95\% confidence an upper limit of 4 yr\(^{-1} \) for pulses of duration of 1 ms and rms amplitude of \( \langle h^2 \rangle^{1/2} \gtrsim 10^{-17} \) [17]. Another similar analysis with only two detectors but 83.3 days of observation time resulted in an upper limit of 1 event per month with \( \langle h^2 \rangle^{1/2} \gtrsim 3 \cdot 10^{-18} \) at the 90\% confidence level [18].

The most promising technology for the detection of gravitational waves by Earth-based instruments now appears to be large laser interferometers [2], which monitor directly the gravitational wave strain that moves differentially its end mirrors. A number of large laser interferometers are now in an advanced stage of their commissioning, but have at best only achieved sensitivities that had already been reached in the past by other prototype instruments. These prototypes were used to set a number of upper limits: TAMA300 was used to set an upper limit of 0.59/hour (90\% confidence) for binary inspirals with two 1.4 \( M_\odot \) compact objects within 6.2 kpc from the detector [19]. A similar result (0.5/hour within the Galaxy) was also published by the LIGO Project, for data obtained with the 40 meters prototype [20]. Using two small interferometers, [21] have set a limit of 1/60 h on bursts in the 800 Hz - 1250 Hz frequency band that have an rms strain amplitude in excess of 5 \( \cdot 10^{-16} \).

### 1.4 Interferometric Detectors

The fundamental design of the interferometers used to look for gravitational waves is that of the classical Michelson interferometer: a laser beam is incident on a beam
splitter, which sends the two resulting beams along two orthogonal paths, until they are reflected by suspended mirrors, and recombine on the beam splitter to form an interference pattern that depends on the distances between the beam splitter and the two mirrors. The mirrors are attached to pendulums which have an oscillation frequency ($\sim 1$ Hz) that is much smaller than the lower edge of the frequency band of interest ($\sim 40$ Hz), and they are therefore a very good approximation to freely falling test masses at the high frequencies where the gravitational wave measurements are made. Consequently, if a gravitational wave $h(t)$ with only a $+$ polarization were incident on such an interferometer along the normal to the plane containing its two arms, which are assumed to have a static length $L$, and if the arms were aligned with the polarization principal axes, the difference $\delta$ between the two arm lengths would oscillate according to

$$\delta(t) = L h(t).$$  \hfill (1.19)

In general, the wave would come from an arbitrary position $(\theta, \phi)$ on the sky, would contain a $+$ and a $\times$ polarization ($h_+$ and $h_\times$), and the polarization axes would be rotated by an angle $\psi$ with respect to the constant-$\phi$ plane, so that the length difference would still be described by Eq.(1.19), but with

$$h(t) = F_+(\theta, \phi, \psi) h_+(t) + F_\times(\theta, \phi, \psi) h_\times(t),$$  \hfill (1.20)

where the beam-pattern factors are given by [13]

$$F_+(\theta, \phi, \psi) = \frac{1}{2} (1 + \cos^2 \theta) \cos 2\phi \cos 2\psi - \cos \theta \sin 2\phi \sin 2\psi$$  \hfill (1.21)

and

$$F_\times(\theta, \phi, \psi) = \frac{1}{2} (1 + \cos^2 \theta) \cos 2\phi \sin 2\psi + \cos \theta \sin 2\phi \cos 2\psi.$$  \hfill (1.22)

The interferometer converts the length difference $\delta$ into a variation of the intensity of the light leaking out of the interferometer. This intensity, in turn, is directly related to the difference of phase $\Delta \Phi$ between the beams that recombine on the beam splitter. For a simple Michelson interferometer, $\Delta \Phi(t) = 4\pi \delta(t)/\lambda$, where $\lambda$ is the wavelength.
of the laser light. One way to increase the phase difference for a certain gravitational wave, apart from increasing the arm length or the laser frequency, is to send the light many times down the interferometer arms, since the phase difference for each round trip down the arms gets added in the interferometer output (unless the total travel time becomes of the order of the period of the gravitational wave, which happens for \(\sim 400\) round trips in the LIGO 4 km arms for a 100 Hz wave). Another approach with a similar effect\(^4\) is to add semi-reflective mirrors between the beam splitter and the end mirrors, thus forming a Fabry-Perot cavity in each arm. If the steady state distance between these input mirrors and the end mirrors is precisely an integer multiple of the laser light wavelength, the light field will resonate within the cavities, consequently producing a large power build-up and a larger phase change in reflections from the input mirrors. For input mirrors which reflect a fraction \(R\) of the incident photons and which transmit the rest without losses, and for perfectly reflective end mirrors, the phase difference is

\[
\Delta \Phi(t) = \frac{16\pi \delta(t)}{(1 - R)\lambda},
\]

i.e. is equivalent to the phase difference produced in an interferometer in which light makes \(4/(1 - R)\) round trips down the arms \((R \approx 97\% \text{ in LIGO})\).

In the absence of all mechanical sources of noise, the sensitivity of a Fabry-Perot interferometer would be limited by fluctuations in the light intensity measurements that are inherent to the quantized nature of light, the so-called shot noise. This noise can be reduced by another improvement over the Michelson interferometer design, which consists in a power recycling mirror located between the laser and the beam splitter. By carefully setting the length of the small cavity formed by the recycling mirror and the two input mirrors (the recycling cavity), the laser light can be made resonant within that cavity, building up power, and effectively increasing the amount of light incident on the beam splitter, by “recycling” the light which would otherwise be reflected toward the laser. In order to allow power recycling, the steady state

\(^4\)The Fabry-Perot design is chosen rather than the delay line principally because it allows the use of smaller mirrors, which are simpler to build. The interferometer, however, is then more sensitive to misalignment or high order spatial modes.
lengths of the interferometer arms are arranged so that the unmodulated components of the beams interfere destructively on the beam splitter (this has the additional advantage of making the gravitational wave measurements only second order sensitive to fluctuations of the laser light intensity). What normally leaks of the antisymmetric port of the interferometer is then only the modulation sidebands that are imposed on the laser light in order to measure all the degrees of freedom of the interferometer, and the smaller quantity of light due to imperfect interference, resulting for instance from defects in the mirror figures, or misalignments of the optics. The photons that are incident on the antisymmetric port photodetector form a Poisson random process, and therefore the light intensity is measured with an error that is proportional to the square root of its mean value. Because the control system tries to keep the interferometer on a destructive interference condition by moving differentially the end mirrors, this photocurrent noise gets transformed into displacement noise of the end mirrors, after being filtered by the instrument transfer function. As a result, the spectrum of the displacement noise due to the photocurrent shot noise scales inversely with the square-root of the power incident on the beam splitter, and scales linearly with the frequency for frequencies above the Fabry-Perot cavity pole (see [22] for details). The recycling mirror increases the power in the light incident on the beam splitter, and hence contributes to the reduction of the displacement shot noise.

Figure 1-2 gives a schematic description of the initial LIGO power recycled Fabry-Perot interferometer, including the error signals used to measure the longitudinal degrees of freedom of the cavities in order to keep the interferometer on resonance in the presence of external disturbances. The laser light is phase modulated, and the modulation frequency is chosen so that the sidebands resonate in the recycling cavity, but not in the arm cavities, thus providing enough information to infer the position of the input and end mirrors. The error signals from the inphase and quadrature components obtained at different points are digitally mixed to form four control signals, one for each longitudinal degree of freedom, which are then appropriately filtered and combined to actuate the mirrors using coils and magnets glued on the optics. In this work, the quadrature component of the error signal at the antisymmetric port (ASY)
Figure 1-2: Optical configuration of a power recycled Fabry-Perot interferometer. The signals from the photodetectors (half disks) are demodulated to extract the inphase (I) and quadrature (Q) components, leading to measurements of various degrees of freedom (L±, l±).

is used as a measurement of the differential motion of the end mirrors, and is called the gravitational wave signal\(^5\). The Michelson control signal, which is used to control the \( l_\perp \) degree of freedom, will be used to implement a veto condition in §3.3.2. In addition, the DC levels of the power in the arm cavities that are measured by the photodetector signals TX and TY are used to decide if the interferometer is “locked” based on the magnitude of the power build-up.

In order to assure the good quality of the length measurements performed by the interferometer, the input laser light must have a low noise level; the light beam frequency, intensity, and spatial profile are controlled by a number of subsystems. The frequency of the laser is stabilized by a hierarchy of systems including a temperature controlled cavity in the pre-stabilized laser (PSL) subsystem, a feedback loop from the mode-cleaner, and a feedback loop from the common mode (L+) of the interferometer.

\(^5\)The transfer function from the physical displacement of the mirrors to the quadrature component of the ASY signal is needed to convert the gravitational wave signal into an estimate of the gravitational wave strain. Information about the calibration of the LIGO instruments is available in [23].
Figure 1-3: Designed displacement sensitivity of the LIGO 4km interferometers.

The mode-cleaner, which is a triangular cavity located between the PSL and the interferometer, is also responsible for filtering out all spatial modes in the light, except the fundamental mode (which has a Gaussian profile). The mode-cleaner is also used for stabilizing the beam jitter by defining an optical axis for the interferometer.

Figure 1-3, which was constructed from data in [24], shows the design sensitivity of the LIGO 4 km interferometers. The most important source of noise below 40 Hz is the seismic noise. Between 40 Hz and 150 Hz, it is the displacement noise produced by the thermal excitation of the pendulum degree of freedom of the suspended mirrors, and by the thermal excitation of the normal modes of the mirrors. At higher frequencies, the noise from thermally excited standing waves in the suspension wires of the optics ("violin" modes) is concentrated in narrow resonances, and therefore have a negligible effect on the sensitivity to a particular source, unless it happens to have a narrow spectrum coincident with one of the wire resonances; the dominating source of noise above 150 Hz is the shot noise.

The displacement sensitivities of LHO-2k and LLO-4k during E7 were two or three orders of magnitude above their design sensitivities (see §5.1 for spectra), and the noise was very significantly non-Gaussian. This was due to the intermediate state
of commissioning of both interferometers. For instance, due to difficulties in acquiring stable operation of the L1 interferometer in the power recycled Fabry-Perot configuration, the recycling mirror of that instrument was misaligned during E7, so that light was not recycled. Additionally, because not all elements of the various subsystems were commissioned, attenuators were placed in front of most photodetectors, so that the signals measured were small compared to the electronics noise, which was therefore the dominating noise in the data for frequencies above $\sim 1$ kHz. To lower frequencies, a variety of electronic and optical noise sources dominated.
Chapter 2

Transient Sources of Gravitational Radiation

"Un cercle n’est pas absurde, il s’explique très bien par la rotation d’un segment de droite autour d’une de ses extrémités. Mais aussi un cercle n’existe pas. Cette racine [d’arbre], au contraire, existait dans la mesure où je ne pouvais pas l’expliquer."

Jean-Paul Sartre

Examples of sources producing transient signals include compact binary coalescences, and stellar collapses. Each of these two classes of sources can be divided into subclasses. Binary coalescences, for instance, may involve two neutron stars, a neutron star and a black hole, or two black holes; a more detailed discussion is presented in §2.1. Similarly, the stellar collapse population might include signals from core collapse supernovae or accretion induced collapses of white dwarfs, or from instabilities that develop after the collapse in the new-born neutron star, assuming such a star is the end point of the collapse; this is discussed in more detail in §2.2 and §2.3.

2.1 Black Hole Binary Coalescences

Gravitational waves carry energy and angular momentum away from the source; in the case of a binary system, this means that the orbit of the two bodies get more tightly
bounded as time goes on, and that the orbital period decreases correspondingly. After a long enough time, the bodies are so close that they merge and form a single, rapidly spinning body.

For this coalescence process to happen in the LIGO band, both members of the binary have to be compact and relatively light objects. Both conditions come from the orbital frequency at merger; assuming two bodies of size \( R \), it is \( \omega \lesssim \sqrt{GM/R^3} \). A binary with total mass \( M \gtrsim 1500 \, M_\odot \) would never enter the LIGO band, even if \( R = R_{Sch} \). On the other hand, a light binary with \( M \sim 1 \, M_\odot \) would also be below the LIGO band if it had \( R/R_{Sch} \gtrsim 130 \), i.e. if \( R \gtrsim 190 \, \text{km} \). This excludes white dwarf binaries and other less dense objects, and hence only binaries made of neutron stars or non-supermassive black holes should be able to enter the LIGO band near the end of their lives.

Neutron star binaries have the simplest orbits from the point of view of the data analysis: the eccentricity \( e \) of a binary system decays with time due to the emission of gravitational radiation, roughly as \( e/e_0 \sim (a/a_0)^{19/12} \), for \( a \) the orbit semi-major axis [25]. In the standard scenario for neutron star binaries formation [26], any initial eccentricity will have plenty of time to disappear as the orbit shrinks to \( a \sim 300 \, \text{km} \), where it enters the LIGO band. The self-gravity of the stars is strong enough to prevent any interaction between them, except near the very end of their inspiral, which is likely to be slightly above the LIGO band [27]. The direct consequence on the data analysis of these observations is that the signal from a certain binary inspiral can be calculated very precisely (using post-newtonian expansions [28]), and the data can be searched very efficiently for the presence of this signal (using matched filtering). The techniques employed to search for neutron stars inspiral [29] signals are radically different from those used in this work, and will not be discussed any further.

As noted in [30], black hole binaries coalescences might easily be detected at a higher rate than neutron star binaries with LIGO, since the more massive binaries should be visible at greater distances, thus making up for the lower expected rate of coalescences. In fact, the volume \( V \) for which the coalescence of binaries consisting
of two objects of equal mass $M$ can be seen by LIGO scales as [30]

$$V \sim M^{2.5},$$  \hspace{1cm} (2.1)

so that even if binaries with total mass $M \sim 40 \, M_\odot$ coalesce at a rate that is $\sim 800$ times smaller than the rate of neutron star binary coalescences, both types of binaries will give the same detection rate with LIGO. Recent estimates of merger rates in the galaxy give $5 \cdot 10^{-7} - 5 \cdot 10^{-6} \, \text{yr}^{-1}$ for black holes binaries forming by dynamical interaction in dense stellar clusters [31], $10^{-7} - 10^{-5} \, \text{yr}^{-1}$ for black holes binaries forming by stellar evolution (i.e. two supernovae in an initially bound binary), and $10^{-7} - 5 \cdot 10^{-4} \, \text{yr}^{-1}$ for neutron star binaries [32].

The scaling relation Eq.(2.1) assumes that matched filtering can be used for all types of binaries. This assumption is invalid at high masses; for instance, the point mass inspiral calculation at 3PN order $[O((v/c)^6)]$ fails at $v/c \sim 0.3$, corresponding to wave frequencies of $\sim 100 \text{Hz}(20 \, M_\odot/M)$ [30]. For massive binaries, the Post-Newtonian expansion will fail before the system even enters the LIGO band. Since the portion of the inspiral following this failure and the final merger of the binary are not presently known with sufficient precision to allow match filtering, the signal from coalescing black hole binaries must be searched for using sub-optimal techniques that are robust against poor signal modeling (a potentially interesting alternative approach is explored in [34]). Additional complications in the use of matched filtering for black hole binaries include the presence of mis-aligned spins and large eccentricities, which could easily be present in the capture scenario of formation; while a single waveform can be generated with good precision for such systems, the increased size of the parameter space makes it impossible (at this point) to search for all possible signals given any practical computational resources.

While detailed modeling is unavailable for black hole binary coalescences, it is still possible to estimate some of the characteristics of the gravitational wave signal. Three principal regimes are interesting in a binary coalescence, as detailed below:

1. **Inspiral**: This phase corresponds to the initial decay of the orbit, where the
two holes are well-separated, and their motion is described with good accuracy by Post-Newtonian techniques (i.e. $v/c \lesssim 0.3$).

2. **Merger:** By convention, the merger starts when the Post-Newtonian techniques fail to describe the evolution of the binary, because the orbit has become significantly relativistic. It then includes the cycles where the two black holes make contact, and merge their event horizons into a single, highly deformed, rapidly spinning hole.

3. **Ringdown:** The merger phase ends when enough of the deformation of the newly formed black hole has been radiated away, so that the hole can be described by linear perturbations of a Kerr metric. These perturbations will also be radiated away, and the end of the ringdown regime will correspond to the point where the hole has radiated all its “hairs”.

As described in §4.5, it will be important in obtaining an upper limit on the rate of coalescence of black hole binaries to inject waveforms into the analysis pipeline, in order to measure its efficiency for detecting such signals. While no complete description of the waveform from a black hole binary merger is currently known, it is possible to develop an approximate signal that is consistent with the present knowledge of these systems. It will be argued later (§5.3) that the analysis system used here is sufficiently robust to handle such errors in the signal predictions. The model developed in [35, Appendix A] will be used. The approach of these authors is to use analytical models for the inspiral and ringdown portions, with a “bridge” merger waveform designed to preserve continuity, but also to satisfy the general energetic conditions known for black holes mergers.

Specifically, the inspiral to merger transition is chosen to occur at the innermost stable circular orbit, defined by the orbital frequency $f_{\text{ISCO}} = 205 \text{ Hz}(20 \ M_\odot/M)$, according to [36]. The inspiral is only defined at first Post-Newtonian order, with the cross and plus polarizations added in quadrature, after root mean square averaging over source angles. After this averaging, the only degrees of freedom of the inspiral waveform is the mass $M$, and the distance $r$ to the binary. The quasi-normal ringdown
waveform is taken from [33]; its temporal structure is defined by an exponentially damped sinusoid at fixed frequency \( f_{\text{qnr}} \) and damping time-scale \( \tau_{\text{qnr}} \), and its angular dependence is defined by a complex function, which depends on the spin of the hole. Again, angular averaging removes the dependence of the waveform on source angles, and the two polarizations are added in quadrature. The parameters of the sinusoid can be expressed in terms of \( M \); for a rapidly spinning hole (with Kerr parameter \( a = 0.98 \)), \( f_{\text{qnr}} = 1320 \text{ Hz}(20 M_\odot/M) \), and \( \tau_{\text{qnr}} = 3.70/f_{\text{qnr}} \). The amplitude of the wave scales with the mass, and is \( 0.05 GM/c^2 r \) initially.

The merger waveform is defined by expressing the merger phase as a fourth order polynomial, and by fixing its five degrees of freedom using the following conditions: (i) the merger signal contains only frequencies between \( f_{\text{ISCO}} \) and \( f_{\text{qnr}} \), (ii) the total energy in the merger and ringdown signals is \( 0.1Mc^2 \), (iii) the duration of the merger is \( 50GM/c^3 \), and (iv) the frequency of the signal and its first time derivative are continuous at the two transitions between the different phases of the coalescence. Conditions (ii) and (iii) are justified in [33] in arguments that are reproduced below. Basically, condition (ii) can be obtained from an argument involving the amount of angular momentum that has to be radiated in the final stages of the merger of two holes: conservation of angular momentum requires that \( S_1 + S_2 + L_{\text{orb}} = J_{\text{rad}} + S_{\text{final}} \), where \( S_{1,2} \) are the spins of the holes before merger, \( L_{\text{orb}} \) is the orbital angular momentum of the binary before merger, \( S_{\text{final}} \) is the spin of the remnant hole after merger, and \( J_{\text{rad}} \) is the angular momentum radiated to gravitational waves. Assuming optimal orientation \( (S_1, S_2, L_{\text{orb}} \text{ aligned}) \), rapidly spinning holes \( (|S_{1,2}| \sim GM^2/4c) \), and the angular momentum at \( f_{\text{ISCO}} \), one gets \( J_{\text{rad}} \sim 0.4GM^2/c \). If all waves are quadrupolar and of frequencies satisfying condition (i), this gives \( E_{\text{rad}} \sim 0.1Mc^2 \), for \( E_{\text{rad}} \) the energy radiated by gravitational waves. Finally, condition (iii) is derived by assuming that the mean luminosity during the merger is the same as the luminosity at the beginning of the ringdown. Since the latter is known, the duration of the merger can be estimated from the lower bound on the energy as estimated above.

This model thus corresponds to signals that are short \( (\sim 10 \text{ ms}) \), strong \( (h \sim 0.1GM/rc^2 \sim 5 \cdot 10^{-19}[M/10 M_\odot][1 \text{ Mpc}/r]) \), and intrinsically broadband \( (\Delta f \sim \)
Figure 2-1: Theoretical distribution of the fraction per unit mass of compact objects in binaries, reproduced from [37].

$f_{\text{qnr}} = 40 \text{ Hz} \sim \text{kHz}$; for the merger specifically, $\Delta f \sim 1115 \text{ Hz}[20 \, M_{\odot}/M])$. They will therefore appear as vertical strips in time-frequency maps, and their power will be spread over many pixels.

Figure 2-1 illustrates a theoretical distribution obtained by [37] for the mass of individual compact objects that are in a binary with another compact object. This distribution was obtained by combining a single-star evolution code and a binary evolution code, and therefore includes a number of effects, e.g. the effects of supernova kicks on the binary. The peak in the distribution that is below $\sim 2.5M_{\odot}$ corresponds to the neutron star population, and the more interesting peak around $10M_{\odot}$ results from the effect of stellar winds: the stellar wind intensity increases with the mass of the zero-age main sequence (ZAMS) star, thus decreasing the mass of the remnant compact object, so that this mass is a weak function of the ZAMS mass. It should be noted that the model used in [37] does not consider the black hole binary formation channel involving capture, as identified by [31], which may produce binaries with heavier holes.
2.2 Core Collapses

The gravitational collapse of a star is a good potential source of gravitational radiation, provided that the star collapses to a compact object, and that the collapse is not spherically symmetric. The former condition comes from the standard requirement that the waves frequency is in the LIGO band, while the latter comes from the need for changes in the quadrupole (or higher order) moment for the radiation of gravitational waves. Two types of collapses that produce compact remnants will be of interest here: core-collapse supernovae, and accretion induced collapses. Both phenomena can involve non-spherical collapses and radiate gravitational waves under certain circumstances; waves generated in the core collapse and bounce of a core-collapse supernova will be discussed in this section, while waves from long-lasting instabilities in the remnant object, when it is a neutron star, will be discussed in the next section.

Core-collapse supernovae are produced by the collapse of the iron core that is formed by nuclear burning in stars that are more massive than $\sim 8M_\odot$. The iron core is initially supported by electron degeneracy and thermal pressures, but as it grows and its mass becomes $\sim M_\odot$, it becomes unstable to the dissociation of iron nuclei into alpha particles at high temperatures and densities, and the resulting electron capture. Without the support from electron pressure, the core collapses until the equation of state stiffens abruptly because of nuclear forces and neutron degeneracy pressure. This point is reached when the core is $\sim 10$ km in radius, and therefore a tremendous amount of gravitational binding energy is released, of the order of $GM_\odot^2/10 \text{ km} \sim 3 \cdot 10^{46}$ J. This energy is transported by neutrinos through the dense core and its surrounding into the outer layers of the star, which explodes in a supernova. These events are rather frequent; the rate of core-collapse supernovae is estimated to be one per 50-140 years in the Galaxy [38], so that these gravitational wave sources could give high detection rates if the strength of the signal is sufficient to cover a fair number of galaxies.

Gravitational waves can be generated only if the collapse process is non-spherical.
The required deformation could be produced by a number of effects, including the effects of rotation, convection, and neutrino anisotropies [39]. It appears from a number of studies that the amount of energy radiated in core-collapse supernovae do not exceed $10^{-6}M_\odot c^2$, even for the most non-spherical collapses [40]. The frequencies of the predicted signals are generally within the LIGO bands, and the maximum predicted strain amplitudes are smaller then $h \sim 10^{-21}$ for sources at 1 Mpc. These numbers, and the details of the produced waveforms, depend on a number of parameters which effects are generally best studied by numerical simulations.

An interesting study of the hydrodynamics of the collapse for a large number of models have been produced by [41] for Newtonian gravity, and recently refined by [40] to include General Relativistic effects. These simulations were designed to cover a range of values for three important parameters of the collapse: the amount of rotational energy, the angular momentum distribution of the progenitor, and the stiffness of the equation of state of the collapsed object. The iron cores are initially approximated by axisymmetric polytropes with adiabatic index $\Gamma = 4/3$, which are in rotational equilibrium. The collapse is induced by reducing abruptly the adiabatic index to a value $\Gamma_r$ in the range $1.28 \leq \Gamma_r \leq 1.325$ ($\Gamma_r$ is the third parameter varied in the simulations, as mentioned in the list above). The stiffening of the equation of state and the shock physics are handled by an analytic approximation for the equation of state. The effects of neutrino transport were neglected; while they have a critical importance for the supernova dynamics, their influence on the early stages of the core collapse might not be so important.

The General Relativistic results were too recent to be incorporated in this work, so another approximation of importance will be that calculations were performed using Newtonian gravity. No “back-reaction” effects were included, and gravitational waves information was extracted using the quadrupole formula [Eq.(1.9)]. The main differences between relativistic and Newtonian simulations are (i) a reduction in the volume of the parameter space allowing multiple bounces of the core, (ii) an increase in the peak density, (iii) a reduction in the signal maximal amplitude and (iv) an increase in the characteristic frequency of the waves. However, the range of strain
amplitudes and wave frequencies covered by the relativistic and Newtonian models were almost the same, and the amount of energy radiated in gravitational waves, when averaged over all the models, varies by less then a factor of two [40].

Figure 2-2 shows the plus polarization\(^1\) for three typical waveforms obtained by [41], which illustrates much of the dynamics of an axisymmetric core collapse. The collapse proceeds in four distinct phases. During the initial infall phase, which lasts between 30 ms and 100 ms, the density of the core increases rapidly, until it reaches nuclear matter densities. At this point the pressure increases suddenly, and the large inertia of the infalling core brings it past the equilibrium point defined by this new pressure; the result is a reversal of the core motion: the core “bounces” back, forming a shock wave at its outer edge. The bounce is the point where the amplitude of the gravitational radiation is maximal. Gravitation eventually damps the outward motion of the core, which settles down to form the neutron star remnant, possibly after a few more bounces or oscillations, depending of the parameters of the model. Finally, if the neutron star spins rapidly enough, it might develop strong instabilities, which will be considered in the next section. Depending on the model parameters, the details of the collapse vary significantly; [41] classify all the simulated waveforms according to three categories:

- **Type I: Regular Collapse** In this type of collapse, the core proceeds cleanly through the three distinct phases of infall, bounce and ringdown, the latter corresponding to small, damped oscillations of the core about its equilibrium position. The top panel of figure 2-2 shows the gravitational wave signal for this type of collapse; the core bounce coincides with the first large negative excursion of the strain amplitude.

- **Type II: Multiple Bounce** For rapidly rotating cores, centrifugal forces can halt the collapse before the density of the core reaches nuclear densities. The core then bounces multiple times on this centrifugal barrier, producing a short

---

\(^1\)Because of the assumptions of axisymmetry and Newtonian gravity, \(h_x = 0\) for all waveforms simulated by [41].
Figure 2-2: The strain amplitude $h$ for three different core collapse models from [41]. All sources are assumed to be at 10 kpc. From top to bottom, the three panels correspond to a waveform of type I, II and III.
burst of gravitational radiation each time, as shown in the middle panel of figure 2-2.

- **Type III: Rapid Collapse** When the change in the adiabatic index $\Gamma$ used to induce the collapse is large, the core collapses more brutally, resulting in a suppressed bounce peak in a gravitational wave signal of reduced amplitude. See the bottom panel of figure 2-2 for such a waveform.

The work of [41], albeit interesting, gives at best only a qualitative description of the collapse process, parameterized over a space that is likely to be larger than that where Nature really distributes the parameters of iron cores from old massive stars. Nevertheless, such a qualitative description is probably sufficient for setting meaningful upper limits, provided that the detection algorithms used are sufficiently robust against errors in the signal predictions.

### 2.3 Neutron Star Instabilities

Rotating neutron stars are unstable due to a number of phenomena, which are generally classified according to the time-scale for the growth of the instabilities. Secular instabilities are driven by fluid viscosity or by gravitational back reaction, and grow on the relevant dissipative time-scale. Dynamical instabilities, on the other hand, are driven by fluid dynamics and gravity, and grow on the time-scale for a sound wave to cross the star, which is much shorter than the secular time-scale. An example of the former is the r-mode instability [42], which results from a $m = 2$ current mode being de-stabilized by gravitational radiation. The modes might be detectable by advanced LIGO detectors [43], but are not strong enough to be of interest here.

One type of dynamical instability that has attracted a lot of interest in the gravitational wave astronomy community is the bar-mode instability. When the star is spinning rapidly enough, is becomes unstable to a $m = 2$ deformation that elongates it along one axis, making it look like an American football spinning end-over-end. If the mode can be sustained for a long time period, this instability could lead to the
emission of copious amounts of gravitational radiation. For even higher rotational frequencies, the star could be disrupted into a number of large pieces that would merge after a few of orbits [44], thus radiating even more energy. It is unclear however if any star could be spun rapidly enough for this fragmentation instability to develop.

In fact, it is not even clear if the rotational velocity required for the growth of the bar-mode instability is achievable commonly. The threshold for the development of instabilities is usually given in term of the parameter $\beta$, which is the ratio of the potential energy to the kinetic energy of a rotating star. Depending on the distribution of the angular momentum within the star, dynamical instabilities can develop if $\beta > 0.14 - 0.27$, the lower limit being obtained for stars with maximum density occurring off-center [45], and the upper limit corresponding to a rigidly rotating, uniform density star (MacLaurin spheroid).

Some of the most recent numerical simulations of core collapses (e.g. [46]) show that $\sim 1$ s after the explosion, $\beta$ is likely to be large enough for the bar-mode instability to develop, unless the proto-neutron star is much less massive than $1M_\odot$ at that time [43, figure 1]. Such a large $\beta$ requires neutron stars to be born with rotation periods of $\sim 1$ms. Observationally, as argued by [43], there is evidence from pulsar observations that at least some neutron stars are born with rotation periods shorter than $\sim 20$ms. Extrapolating the present day observations of pulsar periods to the distribution of initial rotation periods, [47] find that this distribution is Gaussian, with a mean of 7 ms, and with periods shorter than 1 ms located two standard deviations away from the mean (i.e., they represent $\sim 10%$ of the population). However, it may be that these results are biased toward longer initial rotation periods if the spin evolution is faster than assumed due to stronger instabilities radiating gravitational waves in the new-born stars.

Another type of collapse which is likely to yield stars with spins that are large enough to trigger the bar-mode instability is the accretion induced collapse. In this scenario, a white dwarf is pushed above the Chandrasekhar mass limit by accretion from a companion star or by its merger with another white dwarf, and thus becomes unstable to gravitational collapse. As the star density increases, its tem-
perature raises, although it is somewhat limited by neutrino cooling. Depending on the efficiency of the neutrino cooling, the collapsing star may reach a high enough temperature to ignite nuclear burning; in that case the star explodes in a Type Ia supernova. For a more efficient cooling, this high temperature is never reached, and the star collapses until it reaches a new equilibrium as a neutron star, which could then harbor an unstable bar-mode. An upper limit of one accretion induced collapse in the Galaxy per $10^5$ years can be placed by measuring the levels of neutron-rich isotopes in the Galaxy, as these are produced in copious amounts ($\sim 0.1M_\odot$) by such collapses [48].

The rotation frequency of neutron stars formed by accretion induced collapse depends directly on the rotation frequency of the progenitor white dwarf, although it ends up being much higher, due to angular momentum conservation. White dwarfs that are born non-rotating can be spun up to high rotation rates (period $\sim 10$ s) by accretion [43], although characteristic rotation periods for cataclysmic variables typically lie in the range 200 s - 1200 s [49], with minimum observed values around 30 s [50]. These initial data can be used together with accretion induced collapse simulations to estimate the ratio $\beta$ in new-born neutron stars; the work of [43], for instance, gives small values of $\beta < 0.06$, but the evolution of more rapidly spinning progenitors was found to yield large values of $\beta \sim 0.26$ by [51]. Stability studies of the rapidly spinning models of [51] further showed that some of these models were unstable to long-lived bar-mode instabilities [52].

The strength and shape of the gravitational wave signal produced by a bar-mode instability in a neutron star (born in a core collapse or accretion induced collapse) are not known in detail. However, the model described in [52] can be used to generate estimates of these quantities that are useful to the upper limit analysis to be presented below. This model uses linear perturbations to describe the modes of the star; [52] find that for a source at distance $r$ with spin frequency $f$,

$$h = \frac{32\pi^{5/2}G}{5\sqrt{15}c^4r} f^2 Q,$$

(2.2)
\[ \dot{E} = \frac{1024\pi^7 G}{75c^5} f^6 Q^2 , \] (2.3)

where \( h \) is the RMS average of the strain amplitude over source orientation, sky positions, and polarizations, and \( \dot{E} \) is the power radiated. The mass quadrupole moment is denoted \( Q \), and its value is

\[ Q = \alpha 8 \cdot 10^{38} \text{ kg m}^2 , \] (2.4)

for \( \alpha \) the dimensionless amplitude of the bar-mode. Following the finding by [53, 54] using long term numerical simulations that the bar-mode saturates and is persistent when the density perturbation is of the order of the equilibrium density, it is assumed that \( \alpha \) remains constant for an important number of cycles of the gravitational waves, i.e. until the star has lost enough angular to reenter the stability region for bar-modes.

In order to operate at the sensitivity limit of the analysis system used to derive upper limits, i.e. in order to be conservative about the estimation of the efficiency, the frequency evolution of the signal will be chosen to be the hardest one to detect using time-frequency algorithms [see §3.2]. Combining Eqs.(25)-(27) of the paper presented in Appendix A gives this "worst case" frequency evolution,

\[ \dot{f} = 2\pi f^2 \sqrt{\frac{G\dot{E}}{c^5}} , \] (2.5)

which can be integrated if \( \dot{E} \) is assumed constant to give

\[ \frac{1}{f(t)} = 2\pi \sqrt{\frac{G\dot{E}}{c^5}} t + \frac{1}{f(0)} . \] (2.6)

In the actual signals used to measured the efficiency of the analysis pipeline, the power \( \dot{E} \) is set to the value given by Eq.(2.3) with the frequency equal to the signal frequency at time zero, \( f(0) \), and the amplitude of the mode is chosen to be \( \alpha = 0.1 \). The amplitude of the wave is constant, with the value given by Eq.(2.2). The duration
of the signal is set so that the amount of radiated energy is $10^{-3}M\odot c^2$, which, when translated into the amount of radiated angular momentum, is roughly sufficient to bring the star back into the stability region of the bar-mode. Signals therefore last for $1 \text{ s} - 2 \text{ s}$, for frequencies in the range $450 \text{ Hz} - 650 \text{ Hz}$, and see their frequency drop by $\sim 100 \text{ Hz}$ over that time interval, while the signal remains highly coherent.
Chapter 3

Analysis Pipeline

"Plus une illusion est clairement perçue, plus elle a l’air d’une réalité."

Réjean Ducharme

The analysis system described in this chapter represents the smallest analysis system that is complete, in the sense of being scalable to arbitrarily complicated analyses without profound conceptual modifications. The analysis pipeline will assume four data sources: two strain channels from independent interferometers, and one additional diagnostic channel per interferometer.

The first step in the analysis pipeline is to select time segments where data from all four sources are well-behaved (§3.1). These data, which are discretely sampled time-series, are then searched for short bursts, resulting in a list of events (also called triggers below – the details are presented in §3.2 and §3.3). These event lists are then processed through two levels of coincidence analysis: at the first level, veto gates are applied in order to select events from the strain channels that are not coincident with events in the diagnostic channels. At the second level, the strain channels events that have survived the veto gates are compared, and the coincident events are grouped to form gravitational wave candidates. A graphical description of the analysis pipeline is presented in figure 3-1; details about its implementation are given in §3.4.
Figure 3-1: A flow diagram of the analysis pipeline. Data flow along the arrows, and may be time-series or event lists.
3.1 Data Selection

The quality of the data produced by the interferometers is strongly time dependent. Special conditions were applied on these data prior to any additional processing, as enumerated below:

1. **Lock Condition:** The interferometers are not always "locked" to a fringe in their correct operational configurations. Data taken when they are out of lock can not be used for analysis. The DC intensities of the light transmitted through the end mirrors of both arms of an interferometer have to be above a certain threshold for that interferometer to be considered locked. In addition, a safety margin of 60 s is added after the lock acquisition in order to give time for the instrument to stabilize.

2. **Operational State Condition:** All the data segments used in the analysis were free of changes in the parameters of the interferometers. Also, only segments where the appropriate set of whitening/de-whitening filters (located before and after the digital controls, in order to preserve their dynamical range) were turned on were included in the analysis. As for the lock condition, a 60 s safety margin was added after any change in the interferometer settings.

3. **Power Spectral Density Condition:** Because of changes in the alignment and other factors, the noise character is different in every lock stretch. The data are broken into 360 s long segments\(^1\), and the power spectral density of every segment is compared to a reference spectrum. The reference spectrum is constructed by taking the average of 30 of the "best" spectra, as selected by hand from the noise level in the 100-4000 Hz band. For LHO-2k, the comparison is based on the sum of the logarithm of the power over all frequencies above 200 Hz, and on the sum of the logarithm of the power in the 330-390 Hz band. The former of these numbers is used to get an estimate of the mean noise level. The

\(^{1}\)The segment length is chosen to be 360 s because data will be analyzed in segments of such length by the analysis pipeline, see §3.4.
latter number gives the noise in a band that was found to be a good indicator of bad data segments. The logarithm of the power is used because the actual power spans many orders of magnitude over the whole interferometer bandwidth. For LLO-4k, similar criteria are used, except that the frequency band used for the second criterion is 50-200 Hz.

4. **Veto Condition:** Obviously, the analysis pipeline can only be operational when all sub-subsystems are working. The veto trigger generation (the GIDE Analysis in figure 3-1) uses a code that requires some time to generate thresholds, and therefore the whole analysis is only done on data segments where those thresholds are properly acquired.

### 3.2 The TFCLUSTERS Analysis

The heart of the analysis pipeline is the **TFCLUSTERS** algorithm, a time-frequency based method with clustering analysis that allows the efficient detection of weak signals in Gaussian noise. This section is divided into two subsections: §3.2.1 describes the algorithm, and §3.2.2 contains a discussion of the implementation of the algorithm for the analysis pipeline.

#### 3.2.1 TFCLUSTERS: Algorithm

A lot of research has been done on developing the **TFCLUSTERS** algorithm, and on understanding its characteristics as a burst detector, both from the point of view of its operating characteristics (probability of detection vs. probability of false alarm) and from the point of view of its optimality as a detector, when compared to other algorithms. These results, which are beyond the scope of this thesis, will not be reproduced here, but are detailed in the paper included in Appendix A. Only the most important facts about **TFCLUSTERS** will be presented in this section.

The **TFCLUSTERS** algorithm consists of four steps. Step (i) constructs the time-frequency representation of a given time-series. Specifically, the time-series is seg-
mented into equal and non-overlapping segments of duration $T$. Each segment is
Fourier transformed and squared; the result is a two-dimensional picture, with a
time and a frequency axis, giving an estimate of the instantaneous power in the
time-frequency plane. In this work, $T = 1/8$ s, so the frequency resolution of the
spectrogram is 8 Hz.

Step (ii) transforms the spectrogram representation into a black and white picture,
by applying a threshold on the power; pixels above the threshold are called black pix-
els, others are white pixels. For Gaussian noise (and more complicated distributions,
see §3.2.2), the threshold can be computed as a function of frequency, so that all
pixels in the spectrogram have an equal and independent probability of being black;
this black pixel probability is labeled $p$.

Step (iii) applies thresholds on the size of clusters of black pixels. A cluster is
defined as a set of black pixels sharing at least one “edge” with another black pixel
in the set. The size of a cluster is simply the number of black pixels it contains. If
the size $s_1$ of a cluster $\Gamma_1$ satisfies $s_1 \geq \sigma$, for some size threshold $\sigma$, that cluster
is automatically passed to Step (iv). Otherwise, other clusters that are near $\Gamma_1$
are considered, according to a measure of distance $d$ that is isotropic in the time-
frequency plane. If another cluster $\Gamma_2$ of size $s_2$ satisfies $d(\Gamma_1, \Gamma_2) \leq \delta_{s_1, s_2}$, where
$\delta_{s_1, s_2}$ is a distance threshold that depends on the size of the two clusters, then $\Gamma_1$ and
$\Gamma_2$ are linked together to form a generalized cluster, which then also makes it to Step
(iv).

Step (iv) finally applies a threshold on the sum of the power over all pixels in the
clusters or generalized clusters that have survived Step (iii). The threshold is size
dependent, and is designed to reject a fraction $\alpha$ of all events in Gaussian noise, the
rejection fraction being the same for clusters of all size. This step adds a number
of complications when the noise is colored, and therefore it will not be used in the
analysis pipeline ($\alpha = 0$).
3.2.2 TFCLUSTERS: Implementation

The TFCLUSTERS algorithm implementation used in this work is based on the LIGO Data Analysis System [55, 56] and is designed to run on parallel computers. The most important details of this implementation are those dealing with the fact that the instrumental noise is not white nor stationary. The former problem is alleviated by using pre-filtering and the spectral modeling of the noise, as described below. The latter problem is handled by processing the data in relatively short segments (in practice, 360 s), with the assumption that the noise is stationary over that interval, in the sense that the data consist of a stationary colored Gaussian background noise, and any number of non-Gaussian signals.

A stationary white Gaussian noise has a vanishing auto-correlation function everywhere but at zero lag, and correspondingly a flat power spectrum. The noise in the interferometers, on the other hand, have a spectrum that spans many order of magnitudes in power. This leads to two principal problems: correlations in the power at different times, and power leakage. The former problem is generally not a very serious source of error in the analysis, because the time resolution used to construct the spectrogram ($T = 1/8$ s) is much longer than the period of the oscillations in the frequency band of interest ($\gtrsim 100$ Hz), so that the relevant correlations that can not be filtered out with a high pass filtered are simply averaged to nearly vanishing values on timescales of $1/8$ s.

On the other hand, the problem of power leakage leads to correlations between pixels along the frequency direction. Since TFCLUSTERS does not use any window when building the spectrogram in order to preserve the statistical independence of the pixels, as much as 20% of the power in one pixel can leak to pixels below and above it in frequency. In white noise, however, the pixels are independent in frequency, so in addition to the high-pass filter discussed above (which is also essential in eliminating power leakage from low frequencies to the detection band) a whitening filter is applied to prepare the data for TFCLUSTERS. The filter is constructed by a least-square fit of the response function of an infinite impulse response filter (i.e., the norm of its transfer
function) to the inverse of the estimated power spectrum of each interferometer. The estimated power spectrum is the same as the one used in the "Power Spectral Density Condition" of §3.1. The order of the filter is kept small in order not to "over-fit" the data, but is still chosen to be large enough to bring the noise spectra to within 5 dB of a white spectrum, except for a few very narrow lines.

These whitening filters are not designed to remove the narrow features in the power spectra of the noise, such as those resulting from electrical interferences at multiples of 60 Hz. Instead, a model of the power, which allows for such interference, is fitted to the data and used for the thresholds computations. For a given segment of data, the spectrogram \( P_{ij}(y) \) gives a set of independent measurements of the power in different frequency bands\(^2\). The power in each band is modeled as the power that results from the sum of a deterministic interference signal with constant phase and amplitude, and a Gaussian noise with an unknown power spectrum\(^3\) \( S \). In symbols, for \( \Re \tilde{y}_j \) and \( \Im \tilde{y}_j \) the real and imaginary parts of the \( j^{th} \) component of the Fourier transform of the data \( y \), the model is

\[
\begin{align*}
\rho(\Re \tilde{y}_j) &= N(a_j, S_j/2), \quad (3.1) \\
\rho(\Im \tilde{y}_j) &= N(b_j, S_j/2), \quad (3.2)
\end{align*}
\]

where \( a_j \) and \( b_j \) are the constant real and imaginary parts of the interference signal, and \( N(m, \sigma^2) \) is the Gaussian density with mean \( m \) and variance \( \sigma^2 \). The probability density function \( p_{\text{Rice}} \) for the power in the spectrogram is then given by the Rice distribution [57]:

\[
p_{\text{Rice}}(P_{ij}) = \frac{1}{S_j} \exp \left( -\frac{P_{ij} + Q_j}{S_j} \right) I_0 \left( \frac{2\sqrt{P_{ij}Q_j}}{S_j} \right), \quad (3.3)
\]

where \( I_0 \) is the modified Bessel function of the first kind of order zero, and \( Q_j = a_j^2 + b_j^2 \) is the interference signal power. In order to obtain a black pixel probability \( p \) that is

\(^2\)The same notation as in the TFCLUSTERS paper of Appendix A is used here for the spectrogram of the data \( y \), with \( i \) (\( j \)) being the time (frequency) index.

\(^3\)As usual, the power spectrum is defined by \( S = \text{diag}(E[\hat{n}\hat{n}^T]) \), for \( n \) the noise.
uniform in frequency, the following integral equation has to be solved for all values of 
\( j \):

\[
p = \int_{\eta_j}^\infty p_{\text{Rice}}(P_{ij}) dP_{ij},
\]  

(3.4)

where \( \eta_j \) is the threshold to be used in Step (ii) of TFCLUSTERS.

Solving Eq.(3.4) obviously requires the estimation of the power spectrum \( S \) and 
of the signal power \( Q \) for all values of the frequency index \( j \). While a maximum 
likelihood estimation strategy is desirable because it gives the optimal estimator 
assuming one exists, it is not simply applicable here, because the maximum likelihood 
estimator cannot be solved for analytically. Instead, by starting from the moments 
of the distributions described by Eq.(3.3), i.e. from

\[
E[P_{ij}] = S_j + Q_j
\]  

(3.5)

and

\[
\text{var}[P_{ij}] = S_j(S_j + 2Q_j),
\]  

(3.6)

one can estimate the mean and the variance of the power in the data, and solve for \( S \) 
and \( Q \). Figure 3-2 illustrates the magnitude of the error on the black pixel probability 
introduced by using these estimators. The error is small (\( \lesssim 5\% \) near \( Q_j/S_j = 0 \)),
and should not affect other results significantly. However, since the estimators lead to 
values of the black pixel probability that are systematically too large for small values
of the signal power, better estimators must exist.

Finally, large non-Gaussian transients are excluded from the data used for the 
fit to the Rice distribution in order to reduce their bias of the computed thresholds 
toward larger values, which reduces the actual sensitivity of the search. This is done 
by processing the data twice: the first estimation of \( S \) and \( Q \) is done with all the 
available data. A second estimation of the same quantities is then performed, with 
all the power data that are larger than \( Q + \Omega S \) not being used. The value of \( \Omega \) 
that is used is arbitrary, but should be large in order to avoid significant biases on 
the Gaussian component of the noise. In practice, it was found that \( \Omega = 9.5 \) was
Figure 3-2: The value of the black pixel probability as measured from numerical simulations, when the power threshold is computed using Eq. (3.4) for $S$ and $Q$ estimated by the method of moments as described in the text. Each point is the average of 16000 trials, on a sample of size 2880 (i.e., $360 \cdot 8$, which is the number of samples used in the pipeline implementation), where an interference signal with power $Q$ is present.
performing well for the analysis of E7 data. This value was chosen because it was the smallest value of $\Omega$ producing a 1% error on the black pixel probability in Gaussian noise, and was effective at rejecting large transients from the threshold estimation.

### 3.3 Veto Generation

Both vetoes of the analysis pipeline are generated by the same basic procedure, the GIDE algorithm. This algorithm works in the time domain, in order to allow more flexibility with the identification of specific classes of transients, and to cover types of waveforms that are not necessarily detected with high efficiency by time-frequency algorithms. The algorithm works as follow:

1. Acquire a segment of data (typically, 1 s duration).

2. Apply an IIR filter to this segment.

3. Optionally square the filtered segment.

4. Apply a FIR filter to the resulting data.

5. Compute the mean and the standard deviation of the filtered segment auto-regressively. For the mean $m$ and data $x$, $m[i] = W m[i - 1] + (1 - W)x[i]$, where $W = 1/\Lambda + 1/f_s$, for $f_s$ the sampling frequency, and $\Lambda$ the time-scale of the regressive average. An equivalent formula is used for the estimation of the standard deviation, $s[i]$.

6. If the time since the acquisition of the first segment is longer than $\Lambda$, compute the normalized deviation $(x[i] - m[i])/s[i]$, and compare it to a fixed threshold. A stretch of data above threshold form a veto trigger. If the end of a trigger and the beginning of the next trigger are less than a time $t_{\text{min}}$ apart, the two events are merged. The mean and the standard deviation estimates are not updated during triggers.

7. Go back to Step 1.
Typically, the IIR filter is used to select interesting parts of the spectrum, and the FIR filter is used to implement special averaging techniques. The squaring operation is used to look at the power in certain bands, for instance. The regressive estimation of the mean and of the standard deviation is useful in order to follow the non-stationary background noise. In the current analysis, the averaging time-scale was set to $\Lambda = 180$ s, consistent with the choice of the segment length analyzed by TFCLUSTERS (360 s), over which the noise is assumed to be stationary. The threshold on the normalized deviation was generally fairly low, resulting in a fairly large number of triggers. Two additional cuts were applied to the list of triggers generated by running this algorithm on the data; this post-processing of the triggers allowed a relatively rapid optimization of the veto system. The first cut was on the trigger duration. Only triggers with duration longer than 0.2s were retained, corresponding to the observation that shorter triggers showed a negligible correlation with TFCLUSTERS triggers. The second cut was on the maximum of the normalized deviation over all the data points forming a trigger. The optimization of this second threshold is described in §5.2. Details specific to the two different vetoes are described in the following two sections.

### 3.3.1 PSLGlitch

A very important fraction of the non-Gaussian glitches observed in LLO-4k are coincident with transients in the intensity of the light transmitted through the reference cavity of the pre-stabilized laser\(^4\), which are referred to as PSLGlitches. Figure 3-3 shows a segment of data for this channel, containing an episode of glitches.

While they somewhat vary in amplitude, all glitches had a fairly stable waveform shape. The least-square estimation of this common waveform was obtained by a technique using the same tools as the classical method of Principal Component Analysis. First, a large number ($N = 1000$) of glitches in the transmitted light channel were identified by a simple threshold crossing algorithm. All glitches were aligned according to their maximum value, and the data vectors of length $M$ centered on the

\(^4\)During E7, the corresponding channel was L1::PSL-TSS.ACTM.F.
glitches were arranged to form the columns of a $M \times N$ matrix $A$. The singular value decomposition of this matrix was then calculated:

$$A = U \Sigma V^T,$$

(3.7)

where $U$ is a $M \times N$ column-orthogonal matrix, $\Sigma$ is a $M \times N$ diagonal matrix, and $V$ is a $N \times N$ orthogonal matrix. The decomposition is unique up to an orthogonal transformation, and this freedom is used to sort the diagonal elements of $\Sigma$ in increasing order. This decomposition is interesting because of the Eckart-Young theorem [58]: the least-square approximation $\hat{A}$ to a matrix $A$ of rank $Q$, i.e. the matrix $\hat{A}$ that minimizes the trace $\text{tr}[(\hat{A} - A)^T(\hat{A} - A)]$, is given by its truncated singular value decomposition, constructed with only the first $Q$ columns of $U$ and $V$, and the first $Q$ elements of the diagonal of $\Sigma$. Even when the rank of $A$ is unknown, the elements of the diagonal of $\Sigma$ can be examined to determine the importance of the various waveforms, which are described by the columns of $U$: their squared value is the amount of power in $A$ that goes into the corresponding waveform.
Figure 3-4: The square of the first fifty elements of the diagonal of $\Sigma$ (the eigenvalues of the singular value decomposition, labeled $\lambda_i$).

In the present case, the rank of $A$ is expected to be one, i.e. only one waveform type is expected to explain most of the observed power. This is indeed the case, as shown in figure 3-4, where the diagonal elements of the matrix $\Sigma$ are plotted; the first component, illustrated in figure 3-5, explains 52% of the signal power.

The data were pre-filtered with a high pass fifth order elliptical IIR filter with corner frequency at 50 Hz for this analysis. The same filter was used as the IIR filter in the GIDE algorithm, and the first component of the singular value decomposition was used as the FIR filter, effectively implementing a time domain match filter, since the noise spectrum of the light intensity channel was rather flat. The threshold on the normalized deviation was chosen to be 3, and the merging parameter $t_{\text{min}}$ (defined in item 6 of §3.3) was chosen to be 50 ms, since most triggers were quite short.

3.3.2 MICHGlitch

No glitch mechanism was dominating the LHO-2k data as clearly as PSL.Glitch for the LLO-4k data. Nevertheless, a number of large transients were easily observed in the LHO-2k data, and there was in many cases a clear correlation between these
Figure 3-5: The first component of the singular value decomposition, corresponding to the largest eigenvalue.

transients and other control channels of the LHO-2k interferometer. Glitches in the short Michelson interferometer control signal, termed MICHGlitches, were chosen to implement the veto condition on the LHO-2k data stream. This signal defines the differential actuation of the folding mirrors, and therefore responds like the gravitational wave channel to an incident wave, except that its change in response to a wave of a certain strain amplitude is smaller by the ratio of the length of the short Michelson interferometer arm to the length of the full interferometer arm. This ratio is approximately 9:2000, and therefore for any reasonable strain amplitude the short Michelson interferometer control signal can be used as a veto for the gravitational wave channel.

It was impossible to find a small number of reproducible waveforms to describe the MICHGlitches, so a band-limited power monitor was used instead. The IIR filter in the GIDE algorithm was chosen to be a eighth order highpass elliptical IIR filter, with corner frequencies at 60 Hz. The data were squared after the IIR filter, and the result were exponentially averaged by the FIR filter, with a time-scale of 55.5 ms. The merging parameter was chosen to be $t_{\min} = 0.25 \text{ s}$, and the threshold on the
normalized deviation was 2.

3.4 Pipeline Implementation

The only elements of the pipeline itself where a significant amount of design flexibility is available are the coincidence and anti-coincidence gates. They correspond to the same basic operation: given two lists of events, a main and a test list, use the test list to partition the main list into two disjoint subsets, the coincident and the anti-coincident events.

The two anti-coincidence gates used for cleaning the GW channels using the veto triggers operate the same way: the coincidence condition is that the GW event and the veto event overlap in time. The coincidence gate that compares events from the two GW channels takes advantage of the availability of frequency information to define a more stringent condition: the coincidence condition is that the two GW events overlap both in time and in frequency. This is consistent with the expectation that the signal waveforms in LHO-2k and LLO-4k produced by a gravitational wave would be very similar, and that the delay between the two sites will be at most 10 ms, which is much less than the time resolution of TFCLUSTERS.

The last action in the analysis pipeline is the clustering of coincident GW events. This is necessary in order not to overcount events with high multiplicity in the estimation of the rates of events produced by the pipeline. The clustering is implemented by simply merging all events that are less than 1/8 s from each other (so that events at the same time but different frequencies are merged), or that are coincident with the same event from the other GW channel (a situation which indicates that they have the same physical origin).
Chapter 4

Statistics for Upper Limits

"The more to do or to prove, the easier the doing or the proof."

James Joseph Sylvester

The determination of upper limits requires a precise model of the noise in the experiment. This chapter provides a complete description of the model used to analyse the upper limits measurement process (§4.1). In addition, a detailed procedure for the estimation of rate upper limits is developed in §§4.2-4.5, based on this model.

4.1 Pipeline Model

Assuming an abstract model for the analysis pipeline, it will be described in this chapter how to compute an upper limit on the rate of astrophysical events from a certain class \( \mathcal{A} \), with sources located within a certain maximal distance \( \mathcal{D} \) from the Earth. This rate is labeled \( \lambda^*_F \), and is referred to as the foreground astrophysical rate. The foreground instrumental rate \( \lambda_F \) is defined as the rate of coincident events in the analysis pipeline that results from the passage of a gravitational wave from an event of class \( \mathcal{A} \), and is related to the foreground astrophysical rate by

\[
\lambda_F = \epsilon \lambda^*_F, \tag{4.1}
\]
where $\epsilon$ is the efficiency of the pipeline for detecting events from $A$. The efficiency $\epsilon$ is related to the probability $P_D(a, r)$ that an event $a$ from $A$, located at a distance $|r| < \mathcal{D}$, will be detected by the analysis pipeline:

$$\epsilon = \frac{1}{\lambda_F^*} \int_{a \in A} \int_{|r| < \mathcal{D}} p_A(a, r) P_D(a, r) \lambda_F^*(a, r) dadr$$

$$= \frac{E_A\{P_D(a, r) \lambda_F^*(a, r)\}}{E_A\{\lambda_F^*(a, r)\}},$$

where $E_A\{f(a, r)\} \equiv \int_{a \in A} \int_{|r| < \mathcal{D}} p_A(a, r) f(a, r) dadr$. Here, $p_A(a, r)$ is the distribution of sources over event types, positions, and polarization angle (for notational simplicity, the latter is included in the definition of the vector $r$).

It is assumed that the occurrence of astrophysical events in the sphere of radius $\mathcal{D}$ form a Poisson process; the occurrence of coincidences in the analysis pipeline that are due to these events are then also a Poisson process. Of course, what is observed is the sum of this Poisson process and accidental coincidences that have no astrophysical origin. It is assumed that these false coincidences form an independent Poisson process, with background rate $\lambda_B$, and therefore the observed process is also Poisson, with total rate

$$\lambda_{tot} = \lambda_B + \lambda_F.$$

It is assumed that all rates are constant in time, and in general $\lambda_B$ depends on $A$, as the analysis pipeline might be tuned to detect sources of class $A$.

In the actual experiment leading to the measurement of $\lambda_F$, the observable quantity is the number of events in a certain time. The livetime $T$ is defined as the amount of time where a coincidence event with certain parameters could be observed in the pipeline, if it were strong enough. The deadtime is the complementary quantity, i.e. the amount of time where the event could never be observed, for instance because of a veto condition, or because the interferometers were out of lock. The livetime is a function of the properties of the event. For instance, in the two veto gates of the analysis pipeline, bursts of duration $\Delta t$ cannot be detected if their start time is within a time $\Delta t$ of the beginning of a veto event. Hence, the livetime is a function
of $\Delta t$, with $T(\Delta t) \simeq T_{tot} - (\text{number of vetoes})\Delta t$, where $T_{tot}$ is the total observation time, and where veto events are assumed to have negligible durations. For $n(\Delta t)$ the observed number of events of duration $\Delta t$, the rate of events of duration $\Delta t$ is then best estimated by $\hat{\lambda}(\Delta t) = n(\Delta t)/T(\Delta t)$.

The observable total number of events $\mu_{tot}$ produced by the analysis pipeline is the combination of all events with parameters $b \in B$, for $B$ the set of all possible events:

$$\mu_{tot} = \int_{b \in B} p_B(b) \mu_{tot}(b)$$

$$\equiv E_B\{\mu_{tot}(b)\}. \quad (4.5)$$

By definition of the livetime,

$$\mu_{tot}(b) = \lambda_{tot}(b) T[\Delta t(b)], \quad (4.7)$$

and therefore

$$\mu_{tot} = E_B\{\lambda_{tot}(b) T[\Delta t(b)]\}. \quad (4.8)$$

Background and foreground events with parameters $b$ contribute to the total rate of events with parameters $b$ in the same way (i.e., as independent Poisson processes), so $\lambda_{tot}(b) = \lambda_B(b) + \lambda_F(b)$, and Eq.(4.8) can be rewritten as

$$\mu_{tot} = \mu_B + \mu_F, \quad (4.9)$$

where $\mu_B = E_B\{\lambda_B(b) T[\Delta t(b)]\}$ and $\mu_F = E_B\{\lambda_F(b) T[\Delta t(b)]\}$.

The present chapter uses this model of the observations to construct a methodology for computing a confidence interval for $\lambda_F^*$ based on the observables $\mu_{tot}$ and $\mu_B$. The first step is to estimate $\mu_B$ by properly time shifting the data from the two sites, as described in §4.2. Given this estimate, a confidence region for $\mu_B$ and $\mu_F$ can be constructed (§4.3). This region is then mapped to corresponding rates by estimating the amount of livetime in the experiment (§4.4), and finally the estimated efficiency
\( \epsilon (\S 4.5) \) is used together with Eq. (4.1) to obtain a confidence interval for \( \lambda^*_F \).

### 4.2 Background Estimation

If the signals from class \( \mathcal{A} \) are all expected to produce an instrumental signature with duration shorter than \( \Delta t \), adding a delay \( |\delta| > \Delta t \) to one of the two event lists just before the last coincidence gate of the analysis pipeline destroys any potentially present signal, and leads to the background estimator

\[
\hat{\mu}_B = \mu^\delta_{\text{tot}}, \tag{4.10}
\]

where \( \mu^\delta_{\text{tot}} \) is the number of coincidences in the experiment, after the shift. Here \( \hat{\mu}_B > \mu_B \) in general, but \( \hat{\mu}_B \to \mu_B \) as \( \mu_F \to 0 \) \(^1\).

Other then this (small) bias introduced by the possible mixing of signal events with noise events, special care has to be taken not to introduce systematic errors in the measurement of \( \mu_B \). Since the events are defined with a relatively large time resolution \( T \), the lag \( \delta \) has to be an integer multiple of \( T \) in order to avoid the underestimation of \( \mu_B \). The range of \( |\delta| \) should also be limited: since the data are broken into a relatively large number of segments, time shifting one of the datasets reduces the amount of overlap between the datasets, again leading to an underestimation of \( \mu_B \).

For a true number of events \( \mu_B = \lambda_B T_{\text{tot}} \), the measured value is

\[
\hat{\mu}_B = \mu^\delta_{\text{tot}} = E_B \{ \lambda_B (b) (T[\Delta t(b)] - 2N|\delta|) \}, \tag{4.11}
\]

for \( N \) the number of disjoint data segments, and for \( |\delta| \) larger than the duration of the longest expected signal event. The relative error on \( \mu_B \), not including the

\(^1\)After coincidence, \( \mu_{\text{tot}} \approx \mu^1_B \mu^2_B \Delta/T + \mu_F \), for \( \mu^1_B \) the background rate in channel \( i \) before coincidence, \( \Delta \) the coincidence gate time window, and \( T \) the livetime. With an artificial delay, \( \mu^\delta_{\text{tot}} \approx (\mu^3_B \mu^2_B + \mu^1_B \mu^2_F + \mu^1_B \mu^2_B) \Delta/T \), for \( \mu_F \) the foreground rate in channel \( i \), and \( \mu^1_F \geq \mu_F \), with equality if the detection efficiency is unity in channel \( i \). Note the absence of the term \( \mu^1_F \mu^2_F \), because the shifted foreground processes are deterministically related and non-coincident. It follows that \( \mu^\delta_{\text{tot}} \leq \mu_{\text{tot}} \) unless \( (\mu^1_B + \mu^2_B) \Delta/T > 1 \), and that \( \mu^\delta_{\text{tot}} \geq \mu^1_B \mu^2_B \Delta/T \), with equality as \( \mu^1_F \to 0 \).
overestimation due to signal and noise mixing as described above, is then
\[
\frac{\mu_B - \hat{\mu}_B}{\mu_B} = \frac{E_B\{2N[\delta]\lambda_B(b)\}}{E_B\{\lambda_B(b)\Delta t(b)\}},
\]
(4.12)
which can be rewritten as
\[
\frac{\mu_B - \hat{\mu}_B}{\mu_B} \approx \frac{2N|\delta|}{\bar{T}_{tot}}
\]  
(4.13)
according to the approximation described in §4.4, where \(\bar{T}_{tot}\) is defined in Eq.(4.21). In this work, the lag will be chosen to be 5 s, \(N = 199\), and \(\bar{T}_{tot} = 60860\) s [Eq.(5.5)], so the relative error on \(\mu_B\) will be \(\sim 3\%\).

Finally, a potentially better estimator of the number of background events can be obtained by averaging \(\mu_{tot}^\delta\) over the lag. In the present work, this strategy will not be used because of the technical difficulties it introduces in the computation of the confidence region, as explained in §4.3. The loss in quality of the upper limit introduced by this is expected to be small, since the number of background events should be small, and hence \(\mu_{tot}^\delta\) should have a small variance.

4.3 Confidence Regions

Two numbers are measured in a given experiment with the analysis pipeline: \(\hat{\mu}_B\), the estimated number of background coincidences, and \(\hat{\mu}_{tot}\), the total number of coincidences. The measurements have to be performed in independent experiments to avoid biases in the statistics. The population foreground and background number of events (\(\mu_F\) and \(\mu_B\), respectively) in the experiment designed to measure the foreground rate are then inferred from these two numbers. The subject of this section is to describe how to map the measurements \(\hat{\mu}_B\) and \(\hat{\mu}_{tot}\) to the physical parameters \(\mu_F\) and \(\mu_B\) in a rigorous manner that quantifies the measurement errors and the dependence between the measurements.

The confidence region \(R\) is a deterministic function that returns a portion of the \(\mu_F, \mu_B\)-plane given the input \((\hat{\mu}_{tot}, \hat{\mu}_B)\). Suppose that in parallel to the actual experiment observing the process with true parameters \((\mu_F, \mu_B)\), a large number of
other experiments were repeated on independent data that were all obtained from the same test Poisson process, described by \((\bar{\mu}_F, \bar{\mu}_B)\). The confidence regions are defined so that a fraction \(\alpha\) of these experiments will contain the point \((\bar{\mu}_F, \bar{\mu}_B)\); in symbols:

\[
P[(\bar{\mu}_F, \bar{\mu}_B) \in \mathcal{R}(\hat{\mu}_{\text{tot}}, \bar{\mu}_B); (\bar{\mu}_F, \bar{\mu}_B)] = \alpha, \tag{4.14}
\]

where \(\mathcal{R}\) is a realization of the random variable obtained by applying the map \(R\) to data from the test Poisson process. The notation \(P[\cdot; (\mu_F, \mu_B)]\) is used to remind the reader that the experiments are done on a test Poisson process with fixed parameters \((\bar{\mu}_F, \bar{\mu}_B)\). The fraction \(\alpha\) is called the confidence level. The confidence region map \(R\) satisfies Eq.(4.14) for any (physically possible) values of \((\bar{\mu}_F, \bar{\mu}_B)\); this is an essential property, as it guarantees that \(R\) satisfies Eq.(4.14) for the real values \((\mu_F, \mu_B)\), which are unknown.

The Neyman construction [59, 60] provides a simple way to construct the confidence region map \(R\). One starts by selecting a point \((\bar{\mu}_F, \bar{\mu}_B)\), and by computing a region \(\hat{\mathcal{R}}(\mu_F, \mu_B)\) of the measurement space (the \(\hat{\mu}_{\text{tot}}, \hat{\mu}_B\)-plane), so that an arbitrary measurement has a probability \(\alpha\) of falling into that region, assuming it is sampled from the process defined by \((\bar{\mu}_F, \bar{\mu}_B)\):

\[
P[(\hat{\mu}_{\text{tot}}, \bar{\mu}_B) \in \hat{\mathcal{R}}(\bar{\mu}_F, \bar{\mu}_B); (\bar{\mu}_F, \bar{\mu}_B)] = \alpha. \tag{4.15}
\]

The process is repeated for every possible value of \((\bar{\mu}_F, \bar{\mu}_B)\). The confidence region \(\mathcal{R}(\hat{\mu}_{\text{tot}}, \bar{\mu}_B)\) is then constructed by taking the union of all the points \((\bar{\mu}_F, \bar{\mu}_B)\) such that \((\hat{\mu}_{\text{tot}}, \bar{\mu}_B) \in \hat{\mathcal{R}}(\bar{\mu}_F, \bar{\mu}_B)\). In other words, the statement \((\bar{\mu}_F, \bar{\mu}_B) \in \mathcal{R}(\hat{\mu}_{\text{tot}}, \bar{\mu}_B)\) in Eq.(4.14) is equivalent to a statement on the membership of \((\hat{\mu}_{\text{tot}}, \bar{\mu}_B)\) to the region \(\hat{\mathcal{R}}(\bar{\mu}_F, \bar{\mu}_B)\). Therefore,

\[
P[(\bar{\mu}_F, \bar{\mu}_B) \in \mathcal{R}(\hat{\mu}_{\text{tot}}, \bar{\mu}_B); (\bar{\mu}_F, \bar{\mu}_B)] = P[(\hat{\mu}_{\text{tot}}, \bar{\mu}_B) \in \hat{\mathcal{R}}(\bar{\mu}_F, \bar{\mu}_B); (\bar{\mu}_F, \bar{\mu}_B)]. \tag{4.16}
\]

The latter probability is equal to \(\alpha\), hence by construction Eq.(4.14) is satisfied for all values of \((\bar{\mu}_F, \bar{\mu}_B)\).
The confidence region can be constructed with any choice of the shapes of the \( \hat{R} \) regions in Eq.(4.15). The particular choice that will be used in this work is that advocated by [60]. Its principal advantages are the absence of bias in satisfying Eq.(4.14) [i.e. proper "coverage"], and the ability to transition from upper limits to confidence intervals smoothly. The probability \( \alpha \) in Eq.(4.15) is achieved by summing small probability elements \( p(\hat{\mu}_{\text{tot}}, \hat{\mu}_B; \tilde{\mu}_F, \tilde{\mu}_B) \) in order of the decreasing ratio \( r \), which is defined as

\[
    r = \frac{p(\hat{\mu}_{\text{tot}}, \hat{\mu}_B; \tilde{\mu}_F, \tilde{\mu}_B)}{\max_{\tilde{\mu}_F, \tilde{\mu}_B} p(\hat{\mu}_{\text{tot}}, \hat{\mu}_B; \tilde{\mu}_F, \tilde{\mu}_B)},
\]

(4.17)

If Eq.(4.14) cannot be satisfied exactly (in the present application, for instance, \( \hat{\mu}_{\text{tot}} \) and \( \hat{\mu}_B \) have integer values), then a conservative approach is taken, and the probability elements are added until \( \alpha \) is achieved or exceeded.

It is assumed here that the two experiments leading to the measurement of \( \mu_B \) and \( \mu_F \) are independent, so the likelihood function factorizes into a product of probability density functions, one for each experiment:

\[
    p(\hat{\mu}_{\text{tot}}, \hat{\mu}_B; \mu_F, \mu_B) = P_o(\hat{\mu}_B; \tau \mu_B)P_o(\hat{\mu}_{\text{tot}}; \mu_F + \mu_B),
\]

(4.18)

where \( P_o \) is the Poisson density,

\[
    P_o(n; \mu) = \frac{\mu^n e^{-\mu}}{n!},
\]

(4.19)

for \( n \) a positive integer. The parameter \( \tau \) is necessary to scale the number of events when the experiments are performed on stretches of data of unequal lengths. Independently of other parameters, \( \tau = \frac{T_B^{\text{obs}}}{T_F^{\text{obs}}} \), where \( T_B^{\text{obs}} \) (\( T_F^{\text{obs}} \)) is the total observation time for the experiment used to measure the background (foreground).

Elementary calculus shows that the maximum of the probability density of Eq.(4.18) is at \( (\mu_F, \mu_B) = (\hat{\mu}_{\text{tot}} - \hat{\mu}_B / \tau, \tilde{\mu}_B / \tau) \), so that the denominator in Eq.(4.17) is simply:

\[
    \max_{\tilde{\mu}_F, \tilde{\mu}_B} p(\hat{\mu}_{\text{tot}}, \hat{\mu}_B; \tilde{\mu}_F, \tilde{\mu}_B) = \frac{\tilde{\mu}_B \ e^{-\tilde{\mu}_B}}{\tilde{\mu}_B!} \frac{\hat{\mu}_{\text{tot}} \ e^{-\hat{\mu}_{\text{tot}}}}{\hat{\mu}_{\text{tot}}!}.
\]

(4.20)
Figure 4-1: A region constructed from Eq.(4.15) and Eq.(4.18), for $\tau = 1$, $\mu_B = \mu_F = 100$, and $\alpha = 90\%$. The "background" ("total") axis refers to $\hat{\mu}_B$ ($\hat{\mu}_{tot}$). The white cross identifies the point $(\hat{\mu}_B, \hat{\mu}_{tot}) = (\mu_B, \mu_B + \mu_F)$.

Figure 4-1 shows an example of a region constructed with Eq.(4.15) and Eq.(4.18), for $\tau = 1$, $\mu_B = \mu_F = 100$. Figure 4-2 compares the region of figure 4-1 with data points from a simulated Poisson process.

### 4.4 Livetime Estimation

Given a confidence region in the $\mu_F, \mu_B$-plane, it is necessary to define an appropriate mapping to the $\lambda_F, \lambda_B$-plane, in order to gain access to the physically relevant rates. Defining $\lambda_{tot} \bar{T}_{tot} = \mu_{tot}$, one gets

$$\bar{T}_{tot} = \frac{E_B\{\mu_{tot}(b)\}}{E_B\{\mu_{tot}(b)/\mathcal{T}(b)\}}.$$  

(4.21)

If the data are binned over the events' parameters $b$ so that $p_i = p_B(b)db$, this can be reliably estimated by

$$\bar{T}_{tot} \simeq \frac{\sum_i \mu_{tot,i} p_i}{\sum_i \mu_{tot,i} p_i / T_i}.$$  

(4.22)
Figure 4-2: Comparison of the region plotted in figure 4-1 (large closed curve) to simulated data points. The contours show the fraction of points per bin of the 20 by 20 bins histogram covering the whole plot area, over the range [0.005,0.025] in increments of 0.005. $10^6$ points were used in this simulation; 10180 were outside the acceptance region, giving an experimentally measured confidence of 0.898 ± 0.003
In other words, the livetime depends on the characteristics of the events, but is a
deterministic quantity that can be measured directly from the datasets for any choice
of these characteristics. However, since in setting the upper limit on the number
of foreground events an average is taken over events with various characteristics,
the livetimes for those characteristics have to be averaged similarly, with the proper
weights. The way to do this is described by Eq.(4.22).

With the current pipeline design, the livetime depends only on two quantities
directly, namely the duration of the events in the LHO-2k and in the LLO-4k datasets.
For each choice of the values of these two durations, the livetime $\mathcal{T}$ is obtained
by (i) decreasing the start time of all vetoes in both datasets by the maximum of
the two durations, (ii) taking the union of the resulting veto intervals over both
datasets, and (iii) subtracting the result from (ii) from the total observation time.
The coincident events are then accumulated in a histogram in the event durations.
For $N_{j_1,j_2}$, $j_1, j_2 = 1, 2, ..., $ the number of coincident events formed by an event of
duration $j_1 T_{\text{LLO}-4k}$ in the LLO-4k stream, and one of duration $j_2 T_{\text{LHO}-2k}$ in
the LHO-2k stream, the probability $p_i$ in Eq.(4.22) is estimated by $N_{j_1,j_2}/N$, for $N$
the total number of coincident events. The number of coincidences $\mu_{\text{tot},i}$ is estimated by
$N_{j_1,j_2}$, so the actual implementation of Eq.(4.22) is

$$
\tilde{T}_{\text{tot}} = \frac{\sum_{j_1,j_2} N_{j_1,j_2}^2 / \mathcal{T}(j_1 T_{\text{LLO}-4k}, j_2 T_{\text{LHO}-2k})}{\sum_{j_1,j_2} N_{j_1,j_2}^2 / \mathcal{T}(j_1 T_{\text{LLO}-4k}, j_2 T_{\text{LHO}-2k})}. 
$$

(4.23)

This value of $\tilde{T}_{\text{tot}}$ is then used to scale both $\mu_B$ and $\mu_F$. This is only strictly valid if
the background and foreground events individually have the same distribution $p_b(b)$
as the events from the combined (i.e., foreground plus background) process. This
approximation can be interpreted as the analysis pipeline being optimally designed:
any difference in the background and foreground events distributions can be used to
discriminate astrophysical events from noise, and therefore lead to a better analysis
system. While the analysis pipeline presented here is not necessarily assumed to
possess such optimal properties, the approximation is expected to be reasonably good.

Finally, it should be noted that the number of observed coincidences in E7 might
be too small to construct the $N_{j_1,j_2}$ histogram reliably, so that errors on $\bar{T}_{tot}$ are not negligible. The favored solution to this problem will be to generate "fake" coincident events by drawing random events from the populations of events after vetoes in both LLO-4k and LHO-2k, and then retaining only pairs of events that are coincident in frequency, in order to build the $N_{j_1,j_2}$ histogram. Since the populations of events before the final coincidence in the pipeline are very large, this allows the estimation of $\bar{T}_{tot}$ with high accuracy. The $N_{j_1,j_2}$ histogram so estimated should give an unbiased representation of the true distributions, unless there is a large number of correlated events between LHO-2k and LLO-4k. Whether this is the case can be assessed independently from the confidence region on $\mu_B$ and $\mu_F$.

4.5 Efficiency

The efficiency [Eq.(4.3)] has to be measured using numerical simulations, where known calibrated signals are injected into the noise. The approach that will be taken here is the injection into real data, by opposition to the injection of signals into simulated data. The latter approach is in principle superior, because it provides a potentially infinite amount of data independent from those measured, but it requires an excellent understanding of all the statistical properties of the instrumental noise. Given the exploratory nature of this work, no such understanding is available.

Given a signal $a \in \mathcal{A}$ located at some position $\mathcal{R}$ with respect to an interferometer $\mathcal{I}$, a signal $a'$ is formed by (i) taking the linear combination of the two polarizations of the signal, weighted by the antenna pattern of $\mathcal{I}$, (ii) filtering the resulting time series by the transfer function of $\mathcal{I}$, and (iii) applying to the result the \textsc{tfclusters} pre-processing whitening filters. This signal $a'$ is then directly added to the real data, for normal processing by the \textsc{tfclusters} algorithm and the appropriate veto gate. The same procedure is repeated for the other interferometer. Both signals are injected at the same time in the data stream, since typical light travel times between LLO-4k and LHO-2k ($\sim 10$ ms) are small compared to the time resolution of the search code.

The output of the analysis pipeline with signals injected is then compared to its
output without any injection, for a large number of independent segments of data. The probability $\mathcal{P}_D$ that the signal $a$ at distance $r$ produces a coincident event in the analysis pipeline is then estimated from

$$\hat{Q} = \frac{M_{\text{excess}}}{M_{\text{segments}} - M_{\text{vetoed}}},$$

(4.24)

where $M_{\text{segments}}$ is the number of trial segments, $M_{\text{excess}}$ is the number of segments with an injected signal that have an excess of coincident events over the same segments without signals, and $M_{\text{vetoed}}$ is the number of segments where the signal overlaps with a veto. It is necessary to correct the denominator for the signals that are rejected by the veto system because the efficiency is defined with respect to the livetime of the interferometers. The error on $\hat{Q}$ is estimated from the variance of the measurement,

$$\sigma_{\hat{Q}}^2 = \frac{\mathcal{P}_D[1 - \mathcal{P}_D]}{M_{\text{segments}} - M_{\text{vetoed}}} \approx \frac{\hat{Q}[1 - \hat{Q}]}{M_{\text{segments}} - M_{\text{vetoed}}}.\quad(4.25)$$

In practice, in order to improve the speed of the simulations, more than one signal is added at a time to a given segment, the signals being arranged not to be overlapping. In the absence of vetoes, what would then be measured by Eq.(4.24) is the quantity $\hat{P} = 1 - [1 - \mathcal{P}_D]^N$, for $N$ the number of injected signals per segment. The error on the estimator $\hat{P}$ of the probability $\mathcal{P}_D$ is then approximated by

$$\sigma_{\hat{P}} \approx \frac{d}{d\hat{Q}}\left[1 - (1 - \hat{Q})^{1/N}\right] \sigma_{\hat{Q}} \quad(4.26)$$

$$= \frac{1 - \hat{P}}{1 - \hat{Q} N} \sigma_{\hat{Q}}, \quad(4.27)$$

where $\sigma_{\hat{Q}}$ is estimated from Eq.(4.25). Vetoes complicate matters significantly, as they modulate the value of $N$ in the different segments. The solution employed here is to group segments according to $N' = (\text{number of injected signals} - \text{number of vetoed signals})$, and then to take the average of $\mathcal{P}_D$ for each group, weighted by the number of segments in each group. Error estimates are added in quadrature with the same weighting scheme.
In the remaining of this discussion, it will be assumed for simplicity that the sources are uniformly distributed in space around the detector, up to the maximal distance where they have a negligible probability of detection, and that there is no preferred orientation for the sources. In addition, it is assumed that the population is uniform within the class of sources \( \mathcal{A} \), and that the foreground astrophysical rate is independent of the source and source position. Under these assumptions, Eq.(4.3) can be rewritten as

\[
\epsilon = \frac{\int_{\mathcal{A}} \int_{|r| < \mathcal{D}} P_D(a, r) dr da}{V_A \cdot 2\pi \cdot 4\pi D^3 / 3},
\]

(4.28)

where \( V_A = \int_{\mathcal{A}} da \). The signals positions used in the simulations can be chosen so that the beam-pattern of the detector is \( F_+ (\theta, \phi, \psi) = 1, F_\times (\theta, \phi, \psi) = 0^2 \). Changing the angular position or the polarization angle is then equivalent to relocating the source at a new distance \( R' = R / |F_+ (\theta', \phi', \psi')| \). Assuming that the probability \( P_D \) is measured as a function of the source distance \( R \), Eq.(4.28) can be rewritten as:

\[
\epsilon = \frac{\int_{\mathcal{A}} da \int_{r < \mathcal{D}} dr \int d\theta d\phi d\psi \sin(\theta) r^2 P_D(a, r / |F_+ (\theta, \phi, \psi)|)}{2\pi \cdot 4\pi D^3 V_A / 3}.
\]

(4.29)

Switching the order of the radial and the angular integrations, and changing variables \( (r' = r / |F_+ (\theta, \phi, \psi)|) \) leads to

\[
\epsilon = \frac{3}{8\pi^2 D^3} \int_{\mathcal{A}} \frac{da}{V_A} \int d\theta d\phi d\psi \int_{r' \leq \mathcal{D} / |F_+ (\theta, \phi, \psi)|} dr' d\theta' \sin(\theta') |F_+ (\theta, \phi, \psi)|^3 r'^2 P_D(a, r').
\]

(4.30)

This integral can be compute directly, or it can be approximated conservatively by taking the upper bound of the radial integration to simply be \( r' < \mathcal{D} \), which should be adequate for sources that are not expected to be easily detectable beyond \( \mathcal{D} \) (i.e., \( P_D(a, \mathcal{D}) \ll 1 \)). Performing the angular integral leads to

\[
\epsilon = \frac{3\langle |F_+|^3 \rangle}{8\pi^2 D^3} \int_{r' < \mathcal{D}} dr' r'^2 P_D(r'),
\]

(4.31)

where the numerical value of \( \langle |F_+|^3 \rangle \) is 9.8319 (and \( \langle |F_\times|^3 \rangle = \langle |F_+|^3 \rangle \)). The symbol

\[\text{footnote:} \quad ^2\text{A small error is introduced by assuming that LHO-2k and LLO-4k are coplanar and identically aligned. Also, selecting the cross polarization instead of the plus polarization is equally valid.}\]
$\bar{P}_D(r)$ represents the arithmetic mean of the probabilities of detection at distance $r$ of all the simulated signals from class $\mathcal{A}$.

Signals that have both a non-vanishing plus and a non-vanishing cross polarization cannot be handled with the convenient Eq.(4.30). In that case, the strain measured by the interferometers is a linear combination of the two polarizations, and the probability of detection cannot be easily obtained from a simple rescaling of the distance to the source. All simulations presented in this thesis will assume a single polarization waveform and will therefore rely on Eq.(4.30); this approach is justified given the robustness of the TFCLUSTERS detection algorithm to details of the waveforms to be detected.
Chapter 5

Results From the Engineering Run

"Rien n'est beau que le vrai."

Hermann Minkowski

This chapter presents the results of the application of the tools developed in previous chapters to data from the E7 Engineering Run of the LIGO Project. The initial dataset was divided into two subsets: the first one, the *Playground* dataset, consisted of randomly selected segments of data to be used for the tuning of the analysis pipeline, and the study of its efficiency. The remaining of the data formed the *Science* dataset, which was used for the main analysis. The Science dataset was itself divided into two equal subsets: one of these sets, the *Science Set 1*, was used to measure the foreground, while the other set, the *Science Set 2*, was used to measure the background. The Science Set 2 was analyzed with a delay of 5 seconds added to the LLO-4k stream. A precise definition of the three subsets and of all the computer code used in this work is available online [61].

This chapter is divided into six sections. Section 5.1 describes the general characteristics of the E7 data. Section 5.2 describes the optimization of the various thresholds in the analysis pipeline for the E7 noise in order to obtain the best possible upper limit. The efficiency of the optimized pipeline is then studied in §5.3, leading to the measured upper limits presented in §5.4. The different hypotheses used in deriving these upper limits are discussed in §5.5, and are tested to assess possible systematic
<table>
<thead>
<tr>
<th>Cut</th>
<th>Fraction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Veto, PSLGlitch</td>
<td>14.0%</td>
</tr>
<tr>
<td>Veto, MICHGlitch</td>
<td>14.2%</td>
</tr>
<tr>
<td>Power Spectral Density, LLO-4k</td>
<td>13.3%</td>
</tr>
<tr>
<td>Power Spectral Density, LHO-2k</td>
<td>19.1%</td>
</tr>
</tbody>
</table>

Table 5.1: The fraction of the segments satisfying the Lock and OSC conditions that do not satisfy other cuts. The cuts are defined in §3.1.

errors in the measurements. Finally, the upper limits are interpolated to the design sensitivity of the LIGO interferometers in §5.6, in order to estimate the potential reach of the mature detectors.

## 5.1 Data Selection and Data Characterization

As describe in figure 3-1, the first step in the analysis pipeline is to select the data that are going to be useful for the analysis, through a number of criteria (which are defined in §3.1). Of the 408 hours that E7 lasted, the LLO-4k and LHO-2k interferometers were both locked and in a steady state (i.e. the Lock and Operational State conditions were satisfied) for 111.9 hours (27.4% of E7). Dividing these data into 360 s long segments produced 771 segments, or 77.1 hours of data (18.9% of E7). The effect of the other cuts was to throw away almost half of the remaining data: from the initial 771 segments, 425 segments survived these cuts (42.5 hours, 10.4% of E7). Table 5.1 describes what fraction of the rejected data were attributable to the different cuts. Of these final 425 segments, random selection assigned 27 to the Playground dataset, and 199 to both the Science Set 1 and the Science Set 2.

Although it is essential to verify that the events produced by the pipeline analysis are sufficiently stationary and Poissonian to support the analysis model, other aspects of the data provide useful information for the optimization of the analysis. Figure 5-1 shows the power spectral density of the strain noise in the LLO-4k interferometer, and 5-2 shows the equivalent for the LHO-2k interferometer. Both spectra were constructed by averaging the spectra build from the 27, 360 s long segments of the
Figure 5-1: The power spectral density of the strain noise in the LLO-4k interferometer.

Playground dataset, and by normalizing the result with the transfer function of the appropriate detector [62]. The choice of the 400 Hz - 3 kHz bandwidth for the analysis was dictated by these spectra; for LHO-2k, this bandwidth corresponds to all frequencies with noise below $\sim 10^{-19} \text{ Hz}^{-1/2}$, while for LLO-4k it corresponds to frequencies with noise below $\sim 3 \cdot 10^{-19} \text{ Hz}^{-1/2}$.

Especially in the context of a burst search, these spectra only give a rather limited description of the noise level. This can be illustrated by looking at the expected number of events detected by TFCLUSTERS, if the noise were a purely colored Gaussian noise with a few deterministic lines. Using the formalism developed in Appendix A, and the TFCLUSTERS parameters described in table 5.2 for the configuration C, the expected number of clusters per 360 s segment is 0.0065 (± 1%). This represents an expected number of 0.18 cluster for each interferometer in the Playground dataset, and 1.3 clusters in each of the two Science Sets. These numbers should be compared to the entries from the third and fourth rows of table 5.3; the observed number of events is three to four orders of magnitude above the predictions for Gaussian noise! Stated differently, if the noise were perfectly Gaussian, it would be possible to use a
Figure 5.2: The power spectral density of the strain noise in the LHO-2k interferometer.

much lower threshold in TFCLUSTERS, and still obtain the same number of background events; for the most dramatic example, LLO-4k, the first power threshold could be $\eta = 2.71 \epsilon^2$ instead of $\eta = 4.61 \epsilon^2$ as it was in the example above, thus resulting in a $\sim 30\%$ improvement in the distance at which a given source is detectable. It is not assumed anywhere in this analysis that the noise in the interferometers is Gaussian, only that the events produced by the pipeline form a Poisson process. This hypothesis will be tested in §5.4.

5.2 Pipeline Optimization

The difficulty of the optimization of all the degrees of freedom of the analysis pipeline was somewhat reduced by introducing two levels of mostly orthogonal optimizations, thus bringing the problem to a solvable form. In the innermost optimization, all degrees of freedom but the thresholds on the PSLGlitch veto and on the MICHGlitch veto were fixed. The Playground data were then processed by the full analysis pipeline, with different choices of the values of these two thresholds. For every point
in the threshold space that was so sampled, the number of coincident events \( n_B \) and
the livetime \( \mathcal{T} \) were used to compute the \textit{projected} upper limit:

\[
U = \sum_{k=0}^{\infty} \frac{(\tau n_B)^k e^{-\tau n_B}}{k!} U(k; \tau n_B, \tau \mathcal{T}, \alpha). \tag{5.1}
\]

This number corresponds to the expected value of the upper limit for the Science dataset, assuming that the number of background events and the livetime of the Science dataset can be scaled from the corresponding quantities measured in the Playground dataset, using the ratio \( \tau \) of the observation time in the Playground dataset to the observation time in the Science dataset. The function \( U(k; n_B, \mathcal{T}, \alpha) \) gives the upper limit at confidence level \( \alpha \) on the foreground rate, assuming that \( k \) coincident events were detected in an observation run with livetime \( \mathcal{T} \), for a Poisson process producing \( n_B \) background events on average. Consequently, the projected upper limit \( U \) in Eq.(5.1) is the weighted sum of the upper limit that is obtained for \( k \) detected events, with the weights set by the probability to observed \( k \) events from a Poisson process with a rate scaled from the one measured in the Playground dataset.

The function \( U \) is not computed at the level of sophistication described in chapter 4, in order to keep the amount of computations to a manageable level for the optimization. In particular, it assumes no errors on the background rate, thus reducing the dimensionality of the problem to one; the Neyman construction then simply leads to

\[
U(k; n_B, \mathcal{T}, \alpha) = \frac{z(k, \alpha) - n_B}{\mathcal{T}}, \tag{5.2}
\]

where \( z(k, \alpha) \) is obtained by solving

\[
\Gamma[z(k, \alpha), k] = 1 - \alpha, \tag{5.3}
\]

for \( \Gamma[a, b] \) the incomplete Gamma function\(^1\).

Figures 5-3 to 5-9 present results of this optimization. First, figures 5-3 shows how the livetime varies with the two veto thresholds. The livetime was not computed

\[^1\Gamma[z, k] = \int_0^z t^{k-1} e^{-t} dt / \int_0^\infty t^{k-1} e^{-t} dt\]
<table>
<thead>
<tr>
<th>variable</th>
<th>symbol</th>
<th>value</th>
</tr>
</thead>
<tbody>
<tr>
<td>time resolution</td>
<td>$T$</td>
<td>0.125 s</td>
</tr>
<tr>
<td>minimum frequency</td>
<td>$f_{\text{min}}$</td>
<td>400 Hz</td>
</tr>
<tr>
<td>maximum frequency</td>
<td>$f_{\text{max}}$</td>
<td>3 kHz</td>
</tr>
<tr>
<td>size threshold</td>
<td>$\sigma$</td>
<td>5</td>
</tr>
<tr>
<td>distance thresholds</td>
<td>$\delta$</td>
<td>$[0,0,0,0,0,0,2,3,4,4]$</td>
</tr>
<tr>
<td>parameter for two-pass threshold estimation</td>
<td>$\Omega$</td>
<td>9.5</td>
</tr>
</tbody>
</table>
| black pixel probability        | $p$    | Configuration A: 0.08  
|                                |        | Configuration B: 0.05  
|                                |        | Configuration C: 0.01  |
| second power threshold         | $\alpha$ | 0        |

Table 5.2: The three configurations used for TFCLUSTERS in the optimization of the pipeline. All parameters are identical for the LHO-2k and the LLO-4k branches.

with the full formalism of §4.4, but with a lighter one assuming that all GW bursts were of the same duration (picked to be the mode of their actual distribution in the Playground data). Figures 5-4 to 5-6 then show the number of coincident events in the Playground data as a function of the same thresholds, for the different choices of the thresholds in TFCLUSTERS that are described in table 5.2. Finally, figures 5-7 to 5-9 show the projected upper limit, as computed from Eq.(5.1), with a scaling factor $\tau = 23.5$, and a confidence level of 95%.

The principal things to notice from this analysis are enumerated below:

- The projected upper limit surfaces are relatively flat. While the projected upper limit varies by roughly one order of magnitude between each TFCLUSTERS configuration, it changes by at most a factor of 3 when the veto thresholds are varied. For the TFCLUSTERS configuration A, the vetoes can only improve the upper limit by $\sim 20\%$.

- A consequence of this flatness is the occurrence of large errors on the estimation of the optimal point. Its cause probably is the dominating efficiency of the coincidence procedure between LLO-4k and LHO-2k at rejecting background events.

- The PSLGlitch veto is much more important than the MICHGlitch veto. This
Figure 5-3: The ratio of the total observation time to the livetime, as a function of the thresholds on the PSLGlitch veto and on the MICHGlitch veto.

Figure 5-4: The number of coincident events in the Playground dataset, as a function of the thresholds on the PSLGlitch veto and on the MICHGlitch veto, for the TFCLUSTERS configuration A.
Figure 5-5: Same a figure 5-4, but for the TFCLUSTERS configuration B.

Figure 5-6: Same a figure 5-4, but for the TFCLUSTERS configuration C.
Figure 5-7: The 95% confidence level projected upper limit in Hertz, for the TFCLUSTERS configuration A. The cross identifies the chosen optimal point.

Figure 5-8: Same as figure 5-7, but for the TFCLUSTERS configuration B.
Figure 5-9: Same as figure 5-7, but for the TFCLUSTERS configuration C.

can be seen from the larger gradient in the PSLGlitch threshold direction of the projected upper limit surfaces. The effect is more visible for the TFCLUSTERS configuration C, which applies a higher power threshold to the data.

Figure 5-10 shows a comparison of histograms of the raw power\(^2\) in the events, before and after the vetoes were applied. The optimal settings for the veto thresholds for the TFCLUSTERS configuration C were used for building these histograms. The vetoes behave as expected, i.e. they reduce the importance of the high power “tail” of the distributions. It should be noted that the distributions of power are highly non-Gaussian; however, the distributions after vetoes seem to be closer to a Gaussian distribution in the logarithm of the power\(^3\), especially for the LLO-4k data.

\(^2\)The raw power is in ADC counts per root Hertz, i.e. it is the un-calibrated power.

\(^3\)i.e., a log-normal distribution, which would look like two straight lines in figure 5-10, with origin at (0,0), and slope ±1 for ±Z < 0
Figure 5-10: The effect of the vetoes on the distribution of the raw power in the events. Histograms of the number of events vs. the deviation from the mean of the power are shown. \( Z = (P - \bar{P})^3 / 2 \sigma_p^2 |P - \bar{P}| \) is the signed deviation of the logarithm \( P \) of the power, for \( \bar{P} \) the mean of the logarithm of the power, and \( \sigma_p^2 \) its variance. \( P \) is the normalized number of events per bin. Small dots are for events before vetoing, and large dots are for events after the vetoes were applied. The histograms for events after vetoing are computed for \( Z \) calculated with the values of \( \bar{P} \) and \( \sigma_p^2 \) of the unvetted population, so the power scale is the same for both types of dots. The top plot is for LLO-4k, the bottom one is for LHO-2k.
5.3 Efficiency Measurements

All the measurements of the efficiency of the analysis pipeline were performed with the Playground data as input, using the methodology described in §4.5. A number of waveforms were injected into these data in order to assess the variability of the efficiency of the analysis pipeline with the details of the waveform. For each TFCLUSTERS configuration, the veto thresholds were set to the optimal points identified in the optimization described in the previous section.

Five test waveforms were used in order to study the general properties of the efficiency of the pipeline. The results from these simulations are presented in figure 5-11, which shows plots of the probability of detection ($P_D$) as a function of the amplitude signal-to-noise ratio of the injected signal\(^4\), which scales as the inverse of the distance to the source for a given waveform. The different symbols in these figures correspond to different waveforms:

- **Hyphen**: black hole binary coalescence waveform, for a binary made of two $20M_\odot$ holes (§2.1).

- **Asterisk**: black hole binary coalescence waveform, for a binary made of two $5M_\odot$ holes.

- **X mark**: simulated core collapse waveform from the Zwerger-Müller model A1B3G4 (type I, cf. §2.2).

- **Dot**: simulated core collapse waveform from the Zwerger-Müller model A1B3G1 (type II).

- **Open circle**: waveform from a long-lived bar mode instability in a neutron star (§2.3).

---

\(^4\)The signal-to-noise ratio $\rho$ for a signal $h(t)$ is defined as $\rho = \sqrt{\rho_{\text{LHO-2k}}^2 + \rho_{\text{LLO-4k}}^2}$, where $\rho_{\text{LHO-2k}}$ is the LHO-2k signal-to-noise ratio, defined by $\rho_{\text{LHO-2k}}^2 = \int_{1000\Hz}^{3k\Hz} \int |\tilde{h}(f)|^2 / S_{\text{LHO-2k}}(f), \int$ and similarly for LLO-4k. This corresponds to the mean signal-to-noise ratio that would be obtained using match filtering with two detectors optimally oriented to detect a given source.
Figure 5-11: The probability of detection (POD) in the analysis pipeline as a function of signal-to-noise ratio (SNR), for different injected waveforms (see text for details). From top to bottom, the plots correspond to TFCLUSTERS configurations A, B and C.
For a certain choice of maximal distance $D$ for the search, the probabilities of
detection of figure 5-11 can be translated into a pipeline efficiencies using Eq.(4.30).
Dividing the projected upper limits obtained in the previous section by these effi-
ciencies gives projected upper limits on the foreground astronomical rate. For all five
test waveforms, and all distances $D$ of interest, it was found that the TFCLUSTERS
configuration C gave the smallest upper limit. This configuration will therefore be
used for the analysis of the Science dataset.

Errors on the efficiency measurements have a significant effect on the rate upper
limits, especially for large distance reaches where the efficiencies are small, and
hence scale the instrumental rate by a large factor. It was verified using numerical
simulations that the errors on the measured probabilities of detection, estimated as
explained in §4.5, were small enough not to produce relative errors larger than 10%
on the pipeline efficiency. The errors were also relatively smaller when the efficiencies
were larger: for efficiencies larger than 0.5, relative errors were smaller than 5%. Con-
sequently, the astronomical rate upper limits to be quoted below should be considered
precise at the 5% level for reaches where the efficiency is larger than 0.5, but more
imprecise for larger distances, with errors at the 20-30% level when the efficiency is
below 0.1.

The efficiency analyses to be used below for estimating the rate upper limits for
the three classes of sources investigated in this thesis rely on the robustness of the
pipeline against errors in the model for the signal. This robustness was quantified
using numerical simulations, in the following manner: a waveform corresponding to
the black hole binary coalescence of §2.1, for a mass $M = 10.78 M_\odot$, and a distance of
2.9 kpc, was initially generated. This waveform was then Fourier transformed with a
frequency resolution of 10 Hz, and every Fourier component was rotated by a random
angle, picked from the uniform distribution over the interval $[−\theta,\theta]$. The waveform
was transformed back to the time domain; the new waveform had the same spectrum
as the initial one, but was less coherent. Its probability of detection was measured
in the standard way, by injecting the perturbed waveform a number of time in the
Playground data. This procedure was repeated for a number of different values of $\theta$,
and for ten different realizations of the perturbed waveform for each choice of $\theta$; the
results are presented in figure 5-12.

Within the errors, figure 5-12 shows no degradation in the probability of detection
for randomization angles $\theta \lesssim 0.1\pi$ radian; the probability of detection drops to zero,
however, for $\theta \simeq \pi/2$, where the signal has lost most of its coherence. This can
be interpreted directly as a statement on the robustness of the analysis pipeline:
any waveform that has the same power spectrum as the unperturbed black hole
merger waveform but differs from it only by differences in phase will give the same
detection efficiency if the average of the phase differences is less than $\sim 0.1\pi/\sqrt{12}$,
when computed over frequency bins of size 10 Hz. Similar conclusions are expected
to hold for other reference waveforms, and for small differences in the power spectra.

5.4 Results

Table 5.3 gives the number of events of various types in the three datasets. Based
on the number of coincidences after the clustering for the Science Sets 1 and 2, it is
<table>
<thead>
<tr>
<th>After TFCLUSTERS, LHO-2k</th>
<th>Playground</th>
<th>Science Set 1 (foreground)</th>
<th>Science Set 2 (background)</th>
</tr>
</thead>
<tbody>
<tr>
<td>After TFCLUSTERS, LLO-4k</td>
<td>16191</td>
<td>119813</td>
<td>116145</td>
</tr>
<tr>
<td>After MICHGlitch veto, LHO-2k</td>
<td>391</td>
<td>2567</td>
<td>2550</td>
</tr>
<tr>
<td>After PSLGlitch veto, LLO-4k</td>
<td>2271</td>
<td>19613</td>
<td>20319</td>
</tr>
<tr>
<td>LHO-2k coincident with LLO-4k</td>
<td>1</td>
<td>20</td>
<td>18</td>
</tr>
<tr>
<td>LLO-4k coincident with LHO-2k</td>
<td>1</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Coincidences after clustering</td>
<td>1</td>
<td>19</td>
<td>18</td>
</tr>
</tbody>
</table>

Table 5.3: The number of events in the three datasets, at different points in the analysis pipeline.

It is straightforward to apply the methods of §4.3 to obtain a confidence region for the number of foreground and background events. The 95% confidence region is shown in figure 5-13. The projection of this region on the $\mu_F$ axis gives an upper limit on the number of foreground events in the analyzed dataset of

$$\mu_F^{95\%} < 16.5. \quad (5.4)$$

The properly averaged livetime (as described in §4.4) was computed to be

$$\bar{T}_{tot} = (60860 \pm 13)s \quad (5.5)$$

by building thirty histograms of the LLO-4k and LHO-2k burst durations from $10^4$ random picks in the LLO-4k and LHO-2k lists of events after the vetoes, and then using Eq (4.22). The number quoted above for the livetime is the mean of the livetime over these histograms, and the error is its standard deviation. The ratio of the livetime to the number of foreground events gives the 95% confidence upper limit on the foreground instrumental rate:

$$\lambda_F^{95\%} < 2.71 \cdot 10^{-4}\mathrm{Hz}. \quad (5.6)$$

94
Figure 5-13: The 95% confidence region based on the analysis of the Science Sets 1 and 2.

The result given in Eq.(5.6) then only needs to be divided by the efficiency, appropriately averaged over a given source population, in order to obtain the upper limit on the foreground astrophysical rate for that particular source population. This procedure is exemplified for three different types of source populations in the following sections.

5.4.1 Neutron Star Instability

As a first specialization of the instrumental upper limit quoted in Eq.(5.6), the numerical model defined in §2.3, with a range of initial spin frequencies $f(0)$, will be used as a representative sample of the class of signals produced by bar-mode instabilities in neutron stars. It will be assumed that $f(0)$ is uniformly distributed between 450 Hz and 650 Hz. Figure 5-14 shows the estimated probability of detection $P_D$, computed for eleven different values of $f(0)$ in the chosen frequency range. An important spread in the distance at which a certain probability of detection is achieved can be observed from these data; for a probability of detection of 0.5, this distance varies by a factor of almost three. Not surprisingly, this distance decreases regularly from signal to signal.
Figure 5-14: The estimated probability of detection $P_D$, as a function of the source distance $r$, for the bar-mode model of section 2.3, for 11 values of $f(0)$ uniformly distributed between 450 Hz and 650 Hz. Each symbol represents one choice of $f(0)$.

with $f(0)$ increasing, reflecting the shape of the power spectra of the noise in LLO-4k and LHO-2k in the 450-650 Hz range.

Eq.(4.30) is then used to transform the data shown in figure 5-14 into an estimation of the pipeline efficiency $\epsilon$ at detecting sources within a sphere of radius $D$ centered on the Earth; figure 5-15 shows the variation of $\epsilon$ with $D$. Finally, dividing the upper limit on the foreground instrumental rate [Eq.(5.6)] by the efficiency gives the upper limit on the astrophysical rate associated with bar-mode instabilities in neutron stars, as a function of the depth of the search; figure 5-16 shows such a curve.

5.4.2 Core Collapses

Even if new-born neutron stars are not unstable to the bar-mode, the signal from their birth in a core collapse supernova could be detected, as discussed in §2.2. An upper limit on the rate of core collapse supernovae can therefore be deduced, assuming the waveform models simulated by Zwerger and Müller [41]. The Zwerger-Müller catalog of waveforms is too large to use directly for the efficiency estimation; however, since
Figure 5-15: The efficiency of the analysis pipeline at detecting signals from the bar-mode instability model, as a function of the maximum distance to the source.

Figure 5-16: The 95% confidence upper limit on the rate of bar-mode instabilities developing in neutron stars within a certain maximum distance from the Earth.
the waveforms can be classified into a small number of classes and therefore share a number of common characteristics, a subset of ten waveforms is used, and is expected to reflect most of the variability of the full catalog. These waveforms\(^5\) are drawn randomly from the 78 Zwerger-Müller waveforms, and therefore cover uniformly the range of rotation profile of the catalog (from solid rotation to extreme differential rotation), its kinetic energy range (in terms of the parameter \(\beta = \text{(kinetic energy)}/(\text{potential energy})\), \(0.24 \leq \beta \leq 4\)), and its range of changes in the adiabatic index of the equation of state. The amount of energy radiated in the ten selected models ranges from \(6 \cdot 10^{-11} M_\odot c^2\) to \(5.1 \cdot 10^{-8} M_\odot c^2\), and the maximum strain amplitude at 1 kpc ranges from \(3.5 \cdot 10^{-21}\) to \(1.6 \cdot 10^{-19}\).

In order to compute an upper limit, it will be assumed that core-collapse events that occur in nature have parameters distributed so that they produce gravitational wave signals that are uniformly distributed over the ten selected waveforms. The scale of the selected waveforms, however, will be adjusted so that the amount of energy radiated in the collapse is equal to a fixed value in the \(10^{-11} - 10^{-7} M_\odot c^2\) range. The result will be an upper limit that is a function of two variables: the amount of energy radiated, and the volume of the universe that is sampled. In other words, the average in the calculation of the efficiency will be only over the morphology of the signals, and not over their intrinsic amplitude.

The measured probabilities of detection are shown as a function of the distance to the source in figure 5-17, for the ten selected waveforms of the Zwerger-Müller catalog, without rescaling of their radiated energy. The large scatter of the points in distance at a fixed value of the probability of detection reflects the scatter in energy radiated by the model waveforms. The probabilities of detection are translated into a rate upper limit using Eq.(4.30), as shown in figure 5-18.

Figure 5-17: The estimated probability of detection $\mathcal{P}_D$, as a function of the source distance $r$, for the core-collapse model of Zwerger and Müller, as described in §2.2, for 10 different models chosen at random. The amount of radiated energy is that calculated by [41], i.e. it is not scaled. Each symbol represents one model.

Figure 5-18: The 95% confidence upper limit on the rate of core collapses that radiate a certain energy in gravitational waves and that are located within a certain maximum distance from the Earth. Labels on contours indicate the value of the rate in Hertz.
5.4.3 Black Hole Binary Coalescences

In order to derive an upper limit on the rate of equal mass black hole binary mergers, the model of §2.1 will be used. It will be assumed that the model correctly predicts the amount of energy radiated for a given merger, and uncertainties in the waveforms will be explicitly accounted for by averaging over perturbations of the initial waveforms. The perturbed waveforms are constructed by reducing the phase coherence of the model waveforms, as it is done in §5.3, with a maximum rotation angle of ±0.1π. The only free parameter of the waveform model is the mass of the holes forming the binary, so in addition to being averaged over imperfections of the model waveform, the probability of detection is averaged over the black hole masses in order to compute the efficiency, using weights derived from the theoretical mass distribution for black holes in binaries of [37], which was plotted in figure 2-1.

In practice, seven different values of the mass were used, between $3M_\odot$ and $14M_\odot$, and for each combination of mass and source distance, three randomly constructed perturbed waveforms were used. Figure 5-19 shows the measured probabilities of detection, averaged over the perturbed waveforms, and figure 5-20 shows the upper limit, as a function of the volume of space sampled.

5.5 Tests of Hypotheses

The validity of the whole analysis relies heavily on the consistency of the three datasets with each other. For very non-stationary noise, the lengths of the datasets could be insufficient to average out the effects of the varying noise background, so that the variations between the datasets are larger than what would be expected for a stationary random process. The methodology detailed in chapter 4 relies on two hypotheses, which will be discussed below:

1. The number of coincidences in the Science Set 1 and in the Science Set 2 are samples from a Poisson distribution. In other words, if the experiment were repeated a large number of times, the number of coincidences in each Science
Figure 5-19: The estimated probability of detection $P_D$, as a function of the source distance $r$, for the binary black hole merger model. The values plotted are averages over the perturbed waveforms (three per point), and the error bars correspond to the errors from the measurements over each of the three realizations of the perturbed waveform, added in quadrature. Each symbol represents one model.

Figure 5-20: The 95% confidence upper limit on the rate of black hole binary mergers, as a function of the distance reach of the search.
Set, as measured for each experiment, would be distributed according to the Poisson density.

2. The events in the Playground dataset are samples from the same distribution as events from the Science Set 1. This is necessary for the efficiency estimations to be unbiased. In addition to the number of events is the two datasets, all other characteristics of the events should be identically distributed.

5.5.1 Robustness of Poisson Model

As described in §4.3, the computation of the confidence region requires a model for the distribution of the number of observed events. The Poisson distribution was used in this work, but it is legitimate to ask how well this model describes reality.

First, the number of events in the LHO-2k and LLO-4k lists before the final coincidence is not Poisson distributed. The time delay between successive events follows an exponential distribution, for delays longer than approximately one second. On shorter time-scales, the events are highly correlated; this is to be expected, and can be handled by a proper use of clustering techniques (see figure 5-21 for an example). While the time delay distribution is compatible with that of a Poisson process, the number of events per 360 s segment does not follow a Poisson distribution, but rather is closer to a log-Normal distribution\(^6\), as illustrated in figures 5-22 and 5-23. This could be interpreted as (clusters of) events being Poisson distributed within the various 360 s segments, but with a rate that is fluctuating strongly from segment to segment, or on similar time-scales.

It should be noted, however, that what matters for the upper limit calculation is not the distribution of the number of events per segment in LLO-4k or LHO-2k, but the distribution of the number of coincident events over the 199 segments of each of the Science Sets. The observed number of coincident events is too small to allow a precise determination of their distribution. For instance, the two central columns of

\(^6\)The random variable \(x\) has a log-Normal distribution if \(\log x\) is normally distributed. The probability density of \(x\) is then \(p(x) = e^{-(\log x - m)^2/2s^2}/sx\sqrt{2\pi}\), the mean of \(x\) is \(e^{m+s^2/2}\), and its variance is \(e^{2m+s^2}(e^{s^2} - 1)\).
Figure 5-21: The fraction of events as a function of the difference in start time between successive events. The dots are for all LLO-4k events in the Science Set 1, after vetoes, while the crosses are for the same data, with events that have been clustered.

Figure 5-22: The number of 360 s segments as a function of the natural logarithm of the number of events in LLO-4k in each segment, after vetoes, for the Science Sets 1 and 2. The thin curve is the best fit of a log-Normal distribution to the histogram, which has a mean in the logarithm of the number of events of 3.73, and a variance of 1.97.
Figure 5.23: Same as figure 5.22, for LHO-2k. The mean of the log-Normal curve is 1.93, and its variance is 1.31.

table 5.4 give the number of observed coincident events per 360 s segment, and the expected number for a Poisson process with mean equal to 37/398, the total number of observed coincident events divided by the number of segments in the Science Sets. There is a hint of an excess over the Poisson distribution for the cases of two and three coincident events per segment, but the numbers are too small to be conclusive.

Although the data for coincident events do not give direct evidence of deviations from the Poisson model, the observed large variability from segment to segment of the pre-coincidence data arouses suspicion about the viability of this model. Limit theorems may in principle be used to argue that the large variability of the LLO-4k and LHO-2k data streams gets “averaged out”, and that the distribution of the number of coincident events converges to a Poisson distribution in the limit that the probability of two events from LLO-4k and LHO-2k being coincident is small, and in the limit of large datasets. However, it will be argued below that datasets consisting of 199 segments are marginally long enough to converge to a Poisson distribution at the present level of variability, but that if the variance of the logarithm of the number
Table 5.4: For different values of the number of coincident events in a 360 s segment (first column), the number of observed segments in the Science Sets (second column), the expected number of segments for a Poisson distribution (third column), and the mean and standard deviation (quoted as the error on the mean) of the product of the number of LLO-4k and LHO-2k events in the 360 s segments (fourth column).

<table>
<thead>
<tr>
<th>number per segment</th>
<th>number of segments</th>
<th>expected number of segments</th>
<th>$N_1N_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>369</td>
<td>364.7</td>
<td>$(3.25 \pm 4.5) \cdot 10^6$</td>
</tr>
<tr>
<td>1</td>
<td>26</td>
<td>33.7</td>
<td>$(6.59 \pm 5.0) \cdot 10^3$</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>1.56</td>
<td>$(1.11 \pm 0.62) \cdot 10^3$</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>0.0481</td>
<td>N/A</td>
</tr>
<tr>
<td>$\geq 4$</td>
<td>0</td>
<td>$1.1 \cdot 10^{-3}$</td>
<td>N/A</td>
</tr>
</tbody>
</table>

of events was smaller by a factor of $\sim 4$, the convergence would be excellent for datasets of the same duration, assuming the same level of sensitivity for the analysis pipeline as it was for the analysis of E7 data presented in this work.

Given the limited amount of valid data available for E7, it was necessary to rely on numerical simulations to assess how rapidly the distribution of the number of coincident events converges to a Poisson distribution, as the variability of the number of events before coincidence is decreased. The numerical procedure was the following:

1. Draw a random value for the number of events in a 360 s segment for LLO-4k and LHO-2k, using the best-fitted log-Normal distributions of figures 5-22 and 5-23. Label these two numbers $N_1$ and $N_2$, respectively.

2. To simulate the action of taking coincidences between LHO-2k and LLO-4k, define the number of coincident events in the 360 s segment as the random number drawn from a Poisson distribution with mean value $\mu = rN_1N_2$, where $r$ is a small, constant number. This model is compatible with the numbers in the last column of table 5.4.

3. Repeat steps 1 and 2 for all segments in the dataset, and sum the Poisson deviates from 2 to give the simulated total number of coincidences in the dataset.

4. Repeat step 3 ten thousand times to construct the distribution of the total
number of coincidences.

5. Repeat the procedure for a number of simulated datasets, which are drawn from log-normal distributions with scaled down variances.

Although this procedure is much simpler than the actual random process producing the data, it is expected to give a good representation of the variability of the latter.

In an experiment where a coincidence window $\Delta$ is used to compare two point Poisson processes with rates $\lambda_1$ and $\lambda_2$ observed for a time $T$, the expected number of coincidences is $(\lambda_1 T)(\lambda_2 T)(\Delta/T)$. By analogy to this, the value of $r$ in the simulations is set to

$$ r = \frac{\left( \frac{\text{number of coincident events in Science Sets}}{\text{number of events after vetoes in LHO-2k}} \right)}{\left( \frac{\text{number of events after vetoes in LLO-4k}}{\text{length of Science Sets}} \right)} \cdot \frac{360}{360} $$

$$ = 7.20 \cdot 10^{-5}. \quad (5.7) $$

Running the simulations with datasets of the size of the individual Science Sets (199 segments) and with the full variability of the actual data produces the histogram for the total number of coincidences that is shown in figure 5-24. With a reduced $\chi^2$ of 9.1 (70 degrees of freedom) for the difference between the histogram from the simulations and the best fitted Poisson distribution, their is a clear systematic difference between the two distributions. In particular, the ratio of the mean to the variance of the number of events, which should be one for a Poisson distribution, is $\sim 3.3$ in the present case.

If the numerical model is correct and the distribution of figure 5-24 truly represents the distribution of the number of coincident events that would be uncovered if it were possible to repeat the E7 experiment a large number of times, the effects would be significant on the values of the upper limits quoted in this work. For instance, if the simulated distribution of figure 5-24 is used to generate a number of fake results for
Figure 5-24: The number of numerical experiments (out of $10^4$ trials) that gave a certain number of coincident events. The thin curve is the best fitted Poisson distribution, and the vertical dotted line is the measured result from the Science Set 1.

the E7 analysis, and if these results are compared to the region of the $\tilde{\mu}_{\text{tot}}, \tilde{\mu}_B$-plane computed for a confidence of 95% ($\hat{R}(\tilde{\mu}_F, \tilde{\mu}_B)$ in the notation of §4.3), one finds that the coverage of that region by the fake results reaches only 69.7%. This significant under-coverage, if real, would undermine the validity of the analysis leading to the upper limit quoted in Eq.(5.4). However, it should be noted that these concerns result from a numerical model, and are not directly supported by data; until more data are acquired to verify the validity of the model, Eq.(5.4) still represents a valid measure of the upper limit on the number of coincident foreground events.

In particular, it does not seem to be possible at this time to use a different distribution than the Poisson distribution to construct the upper limits. If it were found that the limit of the Poisson distribution is not reached, the only alternative would be to generate the distribution of the number of coincident events, parameterized by the number of foreground events, from numerical simulations that apply the coincidence gate between LHO-2k and LLO-4k lists. This would require an excellent knowledge of the characteristics of the noise in the LHO-2k and LLO-4k event lists after the vetoes,
in addition to other complications (e.g. to average over all possible foreground signal shapes in LLO-4k and LHO-2k). Again, it appears that the Poisson based upper limits are the best the can be achieved with the available data.

Assuming that the complete simulation avenue should ideally be avoided, it is then legitimate to ask how the quality of the data should be improved to achieve the convergence to the Poisson distribution. It can be verified that increasing the length of the analyzed datasets is of no help in this endeavor; the current datasets are long enough that the distribution of the total number of coincident events has converged to an almost normal distribution, with variance larger than the mean by the factor of 3.3 quoted above. Adding segments in the datasets scales the mean and the variance by the same factor, so that their ratio remains unchanged and above the Poisson limit of one. It should be noted, however, that it is possible for the distribution of the number of coincident events to converge to a Poisson distribution if the variability of the number of events in the LHO-2k and LLO-4k streams is reduced by some small factor. This is tested with the current numerical simulations by scaling the variance of the logarithm of the number of events before coincidence in LHO-2k and LLO-4k by the same factor, while keeping the means of these two quantities constant. This procedure progressively transforms the log-normal distributions into better behaved ones as the scaling factor of the variances is reduced. Figure 5-25 shows the ratio of the mean to the variance of the number of coincident events as a function of the scaling factor of the variances. This ratio approaches one for small scaling factors, and is below 1.1 (10% error) for scaling factors smaller than \( \sim 1/4 \). A 10% error on that ratio translates into a \( \sim 1 - 2\% \) error on the probability coverage of the upper limits, which is tolerable.

### 5.5.2 Consistency of Datasets

Under the null hypothesis that no signal is present in any of the three datasets, these should all be consistent with a single distribution for the various parameters describing an event. Table 5.5 presents a series of measurements of the compatibility of the distributions of various quantities in the three datasets. This measurement is
based on the difference of histogram shapes based on the $\chi^2$ statistics formed from the number of events $S_i$ and $R_i$ in the two histograms,

$$\chi^2 = \sum_i \frac{\left(\sqrt{S_i/R_i} - \sqrt{R_i/S_i}\right)^2}{R_i + S_i},$$

(5.9)

for $S = \sum S_i$ and $R = \sum R_i$, and all sums are over bins with $R_i > 0, S_i > 0$. The number of degrees of freedom is equal to the number of such bins. In all cases, it can be seen that the distributions are only occasionally consistent, as measured by the $\chi^2$ statistics. This is likely to be a manifestation of the high variability of the noise from segment to segment: given that variability, the amount of good E7 data available is insufficient for the errors between distributions to converge to well-behaved Gaussian variables, thus giving the observed $\chi^2$ values.

The effects that a larger than Poissonian variability in the number of events in each dataset were covered in the previous section. The effects of the kind of excessive variability that is described by table 5.5 would then only be important by potentially
<table>
<thead>
<tr>
<th>Distribution 1</th>
<th>Distribution 2</th>
<th>reduced $\chi^2$</th>
<th>DOF</th>
<th>$P(H_0)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LHO-2k Central frequency</td>
<td>LHO-2k Central frequency</td>
<td>1.09</td>
<td>248</td>
<td>0.074</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Central frequency</td>
<td>LHO-2k Central frequency</td>
<td>1.62</td>
<td>150</td>
<td>$4.6 \cdot 10^{-11}$</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Central frequency</td>
<td>LHO-2k Central frequency</td>
<td>1.48</td>
<td>152</td>
<td>$8.2 \cdot 10^{-8}$</td>
</tr>
<tr>
<td>Science Set 2</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Central frequency</td>
<td>LLO-4k Central frequency</td>
<td>1.11</td>
<td>352</td>
<td>0.022</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Central frequency</td>
<td>LLO-4k Central frequency</td>
<td>1.61</td>
<td>345</td>
<td>0</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Central frequency</td>
<td>LLO-4k Central frequency</td>
<td>1.77</td>
<td>345</td>
<td>0</td>
</tr>
<tr>
<td>Science Set 2</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Bandwidth</td>
<td>LHO-2k Bandwidth</td>
<td>1.24</td>
<td>60</td>
<td>0.041</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Bandwidth</td>
<td>LHO-2k Bandwidth</td>
<td>1.32</td>
<td>24</td>
<td>0.071</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Bandwidth</td>
<td>LHO-2k Bandwidth</td>
<td>1.75</td>
<td>29</td>
<td>$3.8 \cdot 10^{-4}$</td>
</tr>
<tr>
<td>Science Set 2</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Bandwidth</td>
<td>LLO-4k Bandwidth</td>
<td>1.24</td>
<td>147</td>
<td>0.0027</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Bandwidth</td>
<td>LLO-4k Bandwidth</td>
<td>1.30</td>
<td>103</td>
<td>0.0025</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Bandwidth</td>
<td>LLO-4k Bandwidth</td>
<td>1.15</td>
<td>98</td>
<td>0.073</td>
</tr>
<tr>
<td>Science Set 2</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Duration</td>
<td>LHO-2k Duration</td>
<td>1.34</td>
<td>11</td>
<td>0.13</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Duration</td>
<td>LHO-2k Duration</td>
<td>1.68</td>
<td>12</td>
<td>0.02</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Duration</td>
<td>LHO-2k Duration</td>
<td>3.68</td>
<td>11</td>
<td>$1.2 \cdot 10^{-8}$</td>
</tr>
<tr>
<td>Science Set 2</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Duration</td>
<td>LLO-4k Duration</td>
<td>1.23</td>
<td>22</td>
<td>0.14</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Duration</td>
<td>LLO-4k Duration</td>
<td>6.33</td>
<td>21</td>
<td>0</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Duration</td>
<td>LLO-4k Duration</td>
<td>6.1</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>Science Set 2</td>
<td>Playground</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LHO-2k Power</td>
<td>LHO-2k Power</td>
<td>1.91</td>
<td>48</td>
<td>$2.0 \cdot 10^{-7}$</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LLO-4k Power</td>
<td>LLO-4k Power</td>
<td>1.99</td>
<td>46</td>
<td>$5.1 \cdot 10^{-8}$</td>
</tr>
<tr>
<td>Science Set 1</td>
<td>Science Set 2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.5: A comparison of the distribution of various quantities in the three datasets. For any two distributions, the third column gives the value of the $\chi^2$ computed using Eq.(5.9), divided by the number of degrees of freedom (DOF), which is listed in the fourth column. The last column give the significance of the null hypothesis that the two distributions are identical.
introducing biases in the estimates of the astrophysical rate upper limits if the differences between the Science Set 1 and the Playground set, which is used to measure the efficiency, were significant. It was verified that this was not likely to be the case by recomputing the efficiencies measured for the core-collapse waveforms (§5.4.2) using a randomly selected subset of the Science Set 1, rather than the Playground dataset, as it was done initially. The subset was of the same size as the Playground dataset, and the measured efficiencies agreed for both sets within the measurement errors: 75% of the points 1σ error bars overlapped. This number is smaller than the expectation for Gaussian errors (84%), but the errors are not necessarily Gaussian.

5.6 Extrapolation of Results

The goal of this work was to lay down the design of a simple but fully functional analysis pipeline, to implement this design, and to test it on real data. The performances of this prototype pipeline are expected to scale with a number of factors as the instruments commissioning progresses and the sensitivity and stability of the interferometers improve. This section describes how the upper limits measured with the prototype pipeline can be interpolated to the limits that could be achieved by an international network of interferometers in its mature state. Five specific aspects of the data can be improved:

1. **Noise Stability**: As discussed in §5.5.1, the number of events per unit time, averaged over time-scales of ~ 360 s, might have been fluctuating too strongly during E7 to allow the convergence of the distribution of the number of coincident events to a Poisson distribution. Such a convergence is probably required to perform the data analysis for setting upper limits. It will be assumed that in the future, the variance of the logarithm of the number of events in the LLO-4k and LHO-2k interferometers will be reduced by a factor ≥ 4 compared to its E7 value, to allow the convergence to a Poisson distribution.
2. **Noise Gaussianity:** As mentioned in §5.1, there is an enormous excess of events over the expectations for a Gaussian noise, even after vetoes have been applied to significantly reduce the number of events initially detected. It will be assumed that mature LIGO detectors will have Gaussian noise, i.e. that the non-Gaussian part of the noise that is not vetoed will be insignificant. This will allow the use of lower thresholds in the burst detection algorithms, and should allow an improvement by $\sim 30\%$ of the distance at which a canonical source can be detected.

3. **Strain Sensitivity:** The E7 strain sensitivities were a factor of 300-1000 worse than the design sensitivities [63, Appendix B], depending on frequency, for both LHO-2k and LLO-4k. It will be assumed that the distance at which a source can be detected by mature LIGO detectors scales as the ratio of the signal-to-noise ratio computed with the design strain spectral densities to the signal-to-noise ratio computed with the E7 spectral densities. When all sources in a class do not scale in the same way, the average scaling factor will be used.

4. **Additional Detectors:** The Hanford 4 km interferometer (H1) was not used in the prototype pipeline, but will be available in the future. It will be assumed that the veto system will be sufficiently efficient that H1 events after vetoes will be independent of events from LHO-2k. In addition, it will be assumed that the GEO 600 and Virgo interferometers will be used in conjunction with the LIGO Project instruments. The network signal-to-noise ratio\(^7\) will then be used as explained in the previous item to scale the detection distance for a given source.

5. **Longer Observations:** One year of clean data will all interferometers running will be assumed. The E7 upper limits will then be scaled by $\sqrt{T_{tot}/1 \text{ yr}} = 0.044$, where $T_{tot}$ is the E7 livetime, and is given by Eq.(5.5).

\(^7\)The network signal-to-noise ratio is defined as in §5.2, i.e. it is equal to the signal-to-noise ratios of the individual detectors in the network, added in quadrature.
Table 5.6: The E7 and projected upper limits for different classes of sources.

Table 5.6 gives the projected upper limits for the three classes of sources that were considered previously. The quoted E7 upper limit (second column) is calculated at the distance where the efficiency is one half, which is quoted in the third column. This distance is scaled by the factors described in items 2-4 above, giving the projected reach column in table 5.6. The E7 upper limit computed at 50% efficiency is also scaled by the factor resulting from an increased observation time (item 5), and by the increased distance reach, to give the projected upper limit. This scaling neglects the granularity of the distribution of matter on scales larger than the Galaxy, i.e. it assumes that matter is distributed uniformly out to the projected reach. The fifth column of table 5.6, however, gives the projected upper limit per galaxy, obtained by dividing the E7 upper limit scaled by the observation time factor (item 5) by the number of known galaxies within the projected reach.

The projected limits in table 5.6 are interesting, but they reflect the optimization of the analysis pipeline that was done for the E7 data. Since the E7 dataset was short, it was optimal to tolerate a fairly high false rate (\( \sim 1 \) per hour) to preserve the reach of the search. Consequently, the projected limits have quite interesting reaches, at the cost of having less interesting rates per galaxy. A different tradeoff might very well be chosen for the analysis of future data, depending on the results of a pipeline optimization similar to the one presented here (§5.2). To repeat the numbers that were quoted in chapter 2, theory and observation suggest that the rate in the

---

\(^8\)To be precise, only galaxies with visual absolute magnitudes \( M_V \leq -10.5 \) were considered. Inspection of catalogs of the local group shows that only the Magellanic clouds are within 70 kpc of the Milky Way, and that 22 galaxies satisfying the brightness constraint are within 2 Mpc. For the larger sphere of 70 Mpc of radius, the Schechter's luminosity function [64] is integrated to give \( 3 \cdot 10^8 \) galaxies. The Hubble constant value used was \( H = 75 \) km s\(^{-1}\) Mpc\(^{-1}\).
Galaxy of core collapses is within $7 \cdot 10^{-3} - 2 \cdot 10^{-2}$ yr$^{-1}$, the rate of accretion induced collapses is $\lesssim 10^{-5}$ yr$^{-1}$, and the rate of black hole mergers is in the range $10^{-7} - 10^{-5}$ yr$^{-1}$. Manifestly, the most interesting astrophysical information would be obtained by using a slightly higher threshold, i.e. by restricting the reach and by reducing the false rate. However, if the false rate could be reduced by using coincidences with other instruments (e.g. optical telescopes) without altering the thresholds of the analysis pipeline, a very interesting fraction of the universe would be visible in gravitational waves. This, of course, requires the source of the gravitational waves to emit large amounts of electro-magnetic radiation (or perhaps neutrinos), a condition which might not be met for certain systems (black hole binaries, for instance).

Approximated values of the projected upper limits for higher detection thresholds can be calculated, but they are more error prone than the projected upper limits discussed above due to the non-linear nature of the analysis. Using the following scalings,

$$\left( \frac{\text{upper limit}}{\text{false coincidence rate}} \right) \propto \sqrt{\left( \frac{\text{false rate}}{\text{IFO A}} \right) \times \left( \frac{\text{false rate}}{\text{IFO B}} \right)}, \quad (5.10)$$

a reduction by some factor of the false rate in the two (or more) interferometers of the network leads to a reduction by the same factor of the upper limit. Assuming that all the non-vetoed events detected by TFCLUSTERS result from Gaussian noise, and using the formalism developed in the TFCLUSTERS paper (Appendix A), it is possible to translate this reduction in the false rate into an increase in the power threshold applied to the spectrogram. The square root of this increase can then be used to scale down the distance reach. The optimal background rate for setting a 95% upper limit with one year of data can be calculated using Eq.(5.1), and is $\mu_B = 0.02$ yr$^{-1}$, corresponding to an expected upper limit on the number of foreground events of 2.9 per year. Since the projected background used in constructing table 5.6 is 750 yr$^{-1}$, the false alarm rate in each interferometer has to be scaled down by a factor of $3 \cdot 10^{-5}$, corresponding to a reach decreased by a factor of 2.1, if the efficiency is to be
<table>
<thead>
<tr>
<th>Source</th>
<th>Projected, Rescaled Reach (Mpc)</th>
<th>Projected, Rescaled Upper Limit per Galaxy (yr⁻¹)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core collapse (E_{GW} = 10^{-8} M_\odot c^2)</td>
<td>0.04</td>
<td>6</td>
</tr>
<tr>
<td>Bar-mode instability</td>
<td>1</td>
<td>0.6</td>
</tr>
<tr>
<td>Black hole binary merger</td>
<td>40</td>
<td>(2 \cdot 10^{-4})</td>
</tr>
</tbody>
</table>

Table 5.7: The projected, rescaled upper limits for different classes of sources.

maintained at its value of \(1/2\). Table 5.7 shows the projected, rescaled upper limits, which were computed as the projected limits of table 5.6, except for the distance reaches being rescaled to account for higher thresholds. The projected, rescaled upper limits for the bar-mode instability and the black hole binary merger are within an order of magnitude of the theoretical predictions, and it could be that better data and highly optimized analysis pipelines will allow to bridge that gap.
Conclusion

It is important to develop analysis systems that will do justice to the unprecedented sensitivity to be soon achieved by the interferometric detectors that are being commissioned around the world. It is also desirable that the analysis systems that are developed during this commissioning phase be capable of following the ever increasing quality of the available data with as few modifications as possible, hopefully to the point where observations of gravitational waves will be made on a regular basis. The goal of this thesis was to present such a scalable analysis system, and to give examples of its use for setting upper limits for a number of different astronomical systems.

The pipeline that was analyzed only included two gravitational strain data streams, with a single veto channel each. The formalism discussed in chapter 4 for setting upper limits, and the computer programs used to implement this formalism, can both be directly transposed to a pipeline topology with any number of gravitational strain data streams, each with an arbitrary number of veto conditions. The only complication would come from the optimization of the pipeline parameters (cf. chapter 5), which would then have to be performed in a parameter space of higher dimensionality. It should be the case, however, that a division of the optimization into nearly independent sub-optimizations be possible, thus simplifying the overall process.

This generalization to more complex analysis systems is likely to be necessary as the knowledge of the instruments matures, as the instruments become more sensitive to perturbations in their environments, and as new instruments are added to the network of available detectors. As it is shown in §5.6, the extrapolation of the results obtained with E7 data to the mature network of instruments that should be available within two years gives very interesting numbers for the reach of astronomical searches.
For the black hole binary merger model used in this work, for instance, the future network including the LIGO, GEO and VIRGO interferometers should be able to achieve an efficiency of 50% for sources within 70 Mpc, with a false alarm rate giving an upper limit of $3 \cdot 10^{-3}$ per year per galaxy. If the calculated upper limits scale as expected with the thresholds used in the TFCLUSTERS algorithm, it should even be possible with the future network to place a limit that is one order of magnitude above the most favorable theoretical predictions for black hole binary coalescences and for core-collapse supernovae, if the bar-mode of new-born neutron stars is highly unstable.

It is probably safe to speculate that this last order of magnitude in the upper limit, which translates into a improvement by a factor of three in the distance reach if the false alarm rate is kept constant, could be bridged by further improvements to the analysis pipeline. The use of non-gravitational information, in the form of optical follow-up observations, for instance, might in some cases be used to reduce the background rate, while preserving the reach, i.e. without any change to the thresholds of the analysis pipeline. This could allow the measurement of very interesting "hybrid" upper limits. It would be surprising, however, if a more sophisticated veto system were able to improve by a large factor the performances of the analysis pipeline; as it was shown in §5.2, the coincidence between the sites has a much more important effect on the false alarm rate than the vetoes, even for the highly non-Gaussian E7 data. Nevertheless, a number of improvements of the pipeline could have a significant effect.

An example of such improvements is the optimization of the TFCLUSTERS algorithm, by using wavelet bases, or possibly "cross-spectrogram" built from cross-spectra of two interferometers. If the theoretical knowledge about sources improves sufficiently, it might even be possible to replace the general purpose TFCLUSTERS algorithm by more specific burst detectors, for instance like the one proposed by [34]. The amelioration of the coincidence gate between the interferometers, by comparing more information about the events, such as the power, or shape in the time-frequency domain, or by using time-domain cross-correlations, could also lead to better per-
formances. Finally, it is plausible that the use of better statistical models for the description of the data in computing the upper limits, which may have increased dimensionality, so they not only include the number of events, but other information such as the power, bandwidth or time-frequency shape of the events, will allow better results to be obtained.

If some or all of these ameliorations add up to the required order of magnitude improvement in the efficacy of the analysis pipeline, a significant amount of new astrophysics will be learned, either by rejecting theoretical models predicting rates that are too high, or, more interestingly, by directly observing a burst of gravitational waves from an astronomical system. The same pipeline can be used for setting upper limits and for performing detections; in the language of upper limits developed in this thesis, a detection analysis would be obtained by setting the TFCLUSTERS thresholds to a value that is high enough that a single coincidence observed by the pipeline would lead to a confidence region that is not consistent with zero foreground events. In other words, the pipeline will be operated with a very small false alarm rate, because a background of false detections is not tolerable. The most important conceptual change induced by the transition from the upper limit to the detection analysis is then likely to result from the need for the extrapolation of the coincident event statistics to such small false alarm rates, which are obviously not measurable directly.
Appendix A

The TFCLUSTERS Paper

The following is a research paper (to be submitted to Physical Review D) describing the details of the TFCLUSTERS algorithm, with a particular focus on the definition of the algorithm, its optimality for the robust detection of signals, the calculation of its false alarm rate (analytical and numerical), and an example of the computation of its efficiency for a specific class of signal waveforms.
Time-frequency detection algorithm for gravitational wave bursts

Julien Sylvestre
LIGO Project, Massachusetts Institute of Technology,
NW17-161, 175 Albany St.,
Cambridge, MA 02139, USA.
julien@ligo.mit.edu
(Dated: May 11, 2002)

An efficient algorithm is presented for the identification of short bursts of gravitational radiation in the data from broad-band interferometric detectors. The algorithm consists of three steps: pixels of the time-frequency representation of the data that have power above a fixed threshold are first identified. Clusters of such pixels that conform to a set of rules on their size and their proximity to other clusters are formed, and a final threshold is applied on the power integrated over all pixels in such clusters. Formal arguments are given to support the conjecture that this algorithm is very efficient for a wide class of signals. A precise model for the false alarm rate of this algorithm is presented, and it is shown using a number of representative numerical simulations to be accurate at the 1% level for most values of the parameters, with maximal error around 10%.

I. INTRODUCTION

A number of large laser interferometers [1] are approaching sensitivities to gravitational waves in the \( \sim 10 - 1000 \) Hz frequency band that could be sufficient for the detection of astrophysical events [2]. The signals from these events will be buried deep into the instrumental noise, so that unambiguous detections will be possible only with highly efficient data processing algorithms.

The focus of this article will be on transient sources of gravitational radiation, which will be defined as sources that have a relatively short duration (milliseconds to tens of seconds) and a bandwidth which overlaps at least partially with that of the interferometric detectors. A significant number of such transient sources have been theorized, with various levels of sophistication: the inspiral portion of coalescing compact binaries is well-understood (by post-newtonian expansion techniques [3]), as is the ring-down portion if a black hole results from the coalescence [4], but the merger portion is understood at best only qualitatively [5, 6]. As the mass of the binary increases, the signal-to-noise ratio of the merger portion dominates that of the inspiral and ringdown portions; the merger of \( 10M_\odot - 1000M_\odot \) black hole binaries could be visible at large distances (hundreds of megaparsecs), provided that the waveform can be detected with sufficient efficiency [6, 7]. The collapse of the core of massive stars could also produce detectable signals [8]; depending on the type of progenitor, bar modes, r-modes, fragmentation instabilities or black hole ringdowns could be important sources of gravitational radiation. The details of the waveforms for most of these different mechanisms are far from known; some numerical simulations covering parts of the relevant physics are nevertheless available in the literature [9]. Hence, as it can be seen from the preceding examples, the amount of information about the gravitational wave signal from various sources varies considerably, and this variation is obviously reflected in the efficiency of the algorithms that can be designed for each class of sources.

Only minimal assumptions about the signal will be made in this paper, and therefore the principal characteristic of the algorithm to be presented will be its robustness against poor modeling of the expected signal waveform. Stated differently, this algorithm will be moderately sensitive to a very large class of signals, by opposition to being very sensitive to only a few specific signals. It will correspondingly be useful to search for transient sources that do not have waveforms that are precisely predicted, and to characterize the non-Gaussian, transient component of the noise in the instruments.

A. Summary

Depending on the characteristics of the signals to be detected, the optimization of the detection algorithm is carried out according to various criteria. Some of these criteria are discussed in section II, together with algorithms that were presented elsewhere in the literature and that were shown to be optimal according to these various criteria.

In section III is defined the algorithm which is the subject of this article, to be referred to as TFCLUSTERS, which consists in four specific steps:

(i) The data \( y \) are transformed into a time-frequency representation with fixed time and frequency resolutions \( T \) and \( F \), respectively; the instantaneous power at time \( iT \) and frequency \( jF \), estimated from this representation, is labeled \( P_{ij}(y) \).
(ii) A threshold $\eta$ is applied on the power, in order to retain only pixels with $P_{ij}(y) > \eta$.

(iii) Clusters of pixels with power above threshold are formed by grouping pixels sharing a common side; clusters that do not conform to a fixed set of rules on their size or distance to other clusters are discarded.

(iv) A threshold on the sum of $P_{ij}(y) - \eta$ over the surviving clusters is applied. Data segments containing clusters satisfying this threshold lead to the acceptance of the hypothesis that they contain a signal.

In order to understand the operating characteristics of this algorithm, a simplified version [without the clustering analysis, i.e. consisting only of steps (i), (ii) and (iv)] is shown in section IV to maximize for all signals in $\mathbb{R}^N$ the signal-to-noise ratio among all detectors that are based on the estimation of (a lower bound of) the signal power. This simplified algorithm is especially efficient for signals with a sparse representation in the time-frequency domain. Since most signals are expected to have pixels that present a high degree of spatial correlation in the time-frequency domain, the clustering analysis of TFCLUSTERS is an interesting way to capitalize on that property to filter out a large portion of the noise, as it is shown in section V.

An analytical method for computing the false alarm rate associated with TFCLUSTERS is developed, with the details for the clustering analysis being presented in section VA. Using the computer-generated enumeration of all the possible clusters of a certain size that can be formed, it is shown that large clusters are exponentially unlikely to occur when only noise is present. The rate of occurrence of pairs of clusters separated by a certain distance is computed in an analogous manner. Numerical simulations (section VI) confirm that these theoretical predictions are accurate in the regime of operation relevant to TFCLUSTERS, with errors on the predicted false alarm rates around 1% in most situations, and with maximum errors around 10%, due to the exclusion of higher-order terms in the theoretical modeling, or to unmodeled finite size effects.

One example of a complete analysis of the performances of TFCLUSTERS for a short, narrow-band signal, including the optimization of its efficiency at fixed false alarm rate, is presented in section VB. The efficiency of TFCLUSTERS is compared to that of the (unrealistic) ideal power detector (which assumes a knowledge of the signal duration and central frequency) as a function of the signal-to-noise ratio of the signal; at fixed probability of detection, the reduction in signal-to-noise ratio for TFCLUSTERS is consistently less than about 30%.

II. TRANSIENT DETECTION

The transient detection problem consists in choosing between two possibilities: the observed data $y(t)$, $0 \leq t \leq T$, consist of a signal term $s(t)$ and a noise term $n(t)$, where the noise is additive and is assumed to be white Gaussian with zero mean and unit variance [10]:

$$y(t) = s(t) + en(t), \quad (1)$$

or the data are noise alone:

$$y(t) = en(t). \quad (2)$$

In the simplest case, $s(t)$ is fixed and known. Perhaps the most natural optimality criterion is then to maximize the probability of detection when the signal is present [Eq. (1)] for the probability of detection when no signal is present [Eq. (2)] being smaller than some preassigned false alarm probability. This so-called Neyman-Pearson criterion leads to the use of the amplitude $Q$ of the signal, optimally estimated [11] from the scalar product

$$Q = \langle y(t), s(t) \rangle, \quad (3)$$

as the statistic on which the decision between Eq. (1) and Eq. (2) is made.

In the case where $s(t) \in W$ for some function space $W$, one can integrate the signal distribution to reduce the problem to the simple case of discriminating between two fixed probability distributions, as is done when $s(t)$ is fixed and known. This requires the knowledge of the prior distribution $p[s(t)]$, and results in using the likelihood ratio

$$A[y(t)] = \frac{\int_{s(t) \in W} p[y(t)|s(t)]p[s(t)]ds(t)}{p[y(t)|0]} \quad (4)$$

as the detection statistic. When the prior $p[s(t)]$ is unknown, the choice of an optimality criterion is not as simple as for the fixed signal case; the modified Neyman-Pearson criterion maximizes the minimum of the probability of detection over all $s(t) \in W$, for the false alarm probability being below some preassigned value, and is therefore an interesting
"conservative" choice. It is well-known that the optimal algorithm in terms of the modified Neyman-Pearson criterion is obtained by choosing the prior \( p(s(t)) \) that is least-favorable, i.e. the prior that minimizes the minimum over \( W \) of the probability of detection when the signal is present, at fixed false alarm rate [12]. In particular, using the generalized likelihood ratio

\[
\Lambda'(y(t)) = \max_{s(t) \in W} \frac{p[y(t)|s(t)]}{p[y(t)|0]} 
\]

as the detection statistic does not guarantee optimality [13].

For signals with excellent theoretical models, such as binary neutron star or black hole inspirals, the function space \( W \) is compact enough that the least-favorable prior is approximately uniform in \( s(t) \), so that the generalized likelihood ratio [Eq. (5)] derives from Eq. (4) as the optimal detection statistic. This leads to a convenient implementation [14] which simply thresholds on the maximum of the correlation [Eq. (3)] over a filter bank, which is a discretely sampled version of the function space \( W \).

For the case where the signal space is not simple enough to allow match filtering, that is when the signal is not well-modeled, a number of incoherent methods have been proposed in the literature. One of them is the so-called excess power detector [15], which basically threshold on the power integrated over a large number of different shapes at different positions in the time-frequency plane. In their discussion of the optimality of the excess power detector, the authors of [15] use a prior uniform in the “whitened” signal subspace of all waveforms of finite duration and bandwidth, which is imposed from physical arguments, and their result is therefore only a proof of optimality with respect to that particular prior. In particular, this prior is not necessarily least-favorable, and their result is therefore not a proof of optimality in the modified Neyman-Pearson sense. Similarly, the author of [16] chooses a different prior, which is uniform in the “unwhitened” signal space, to derive a detector which is optimal with respect to this prior, and which is similar to the excess power detector, but is perhaps better adapted to colored instrumental noise.

In addition to these detectors, a number of ad hoc methods have been proposed to solve the transient detection problem for unmodeled transients. They are based on the analysis of patterns in the time-frequency plane [17], or on the time-domain analysis of the data with various filters [18]. Only the authors of [18] discuss optimality, from a numerical point of view, i.e. by using a small number of simulated signals from [9] that are injected into noise in numerical simulations in order to compare the performances of a few different detection algorithms.

When the noise is Gaussian, the likelihood ratio in Eq.(4) can be rewritten as [19]:

\[
\Lambda[y(t)] = \exp \left[ \frac{T}{2} y(t) \tilde{s}(t) dt - \frac{1}{2} \int_0^T \tilde{s}^2(t) dt \right],
\]

where \( \tilde{s}(t) \) is the causal minimum mean-square error estimator of \( s(t) \), i.e.,

\[
\tilde{s}(t) = E[s(t)|y(\tau)],
\]

for \( E[y(\tau)] \) the expectation over the noise given the observation \( y(t) \) for \( \tau < t \), assuming the model of Eq.(1). The symbol \( \int \) represents the Itô stochastic integral [20]. Of course, the computation of \( \tilde{s}(t) \) in Eq.(7) requires the knowledge of the prior \( p(s(t)) \), but the structure of Eq.(4) suggests that an efficient approach to problems for which the integration in Eq.(4) cannot be carried out effectively (due to a complex or unknown prior, for instance) might be to develop efficient estimators of the signal, and to use them as filters to test for the presence of signals in the data (i.e., to use an estimator-correlator design).

It is shown in [21] that (i) transforming the (discretely sampled) observations \( \{y_i : i = 1 \ldots N\} \) to a wavelet basis, (ii) applying on the transformed observations \( \{\tilde{y}_i\} \) the non-linearity

\[
\tilde{s}_i = \begin{cases} 
0 & \text{if } |y_i| < \eta \\
\tilde{y}_i - \eta & \text{if } \tilde{y}_i > \eta \\
\tilde{y}_i + \eta & \text{if } \tilde{y}_i < \eta,
\end{cases}
\]

for a threshold \( \eta \sim \epsilon \sqrt{2 \log(N)/N} \), and (iii) transforming the truncated observations \( \{\tilde{s}_i\} \) back to the time domain lead to an estimator \( \{\hat{s}_i\} \) that is optimal in the sense of giving the minimal expected least-square error, the minimum being over all signals in \( \mathbb{R}^N \). Preserving the structure of Eq.(6), a simple quantity to threshold on is

\[
\sum_{i=1}^N \left( y_i \tilde{s}_i - \tilde{s}_i^2/2 \right).
\]

124
which can be rewritten as

$$\sum_{i=1}^{N} \left( \frac{\hat{s}_i^2}{2} + \eta|\hat{s}_i| \right).$$  \hspace{1cm} (10)

The algorithmic structure of **TFCLUSTERS** corresponds to this model; the two principal differences are that a short-time Fourier basis is used instead of a wavelet basis, and that an additional step involving the analysis of the correlations between the non-zero components of the estimator $\hat{s}$ is introduced.

### III. ALGORITHM

It will be assumed that $N$ data corresponding to the gravitational wave strain are read and stored as real numbers into a row vector $y$. The data are time ordered and uniformly sampled with sampling frequency $f_s$, so the $i$th component $y_i$ of the vector $y$ is the value of the measured strain at time $i/f_s$: $y_i = y(i/f_s)$.

Step (i) of **TFCLUSTERS** is the construction of the time-frequency representation of the data from the spectrogram:

$$P_{ij}(y) = |\hat{y}_{ij}|^2,$$  \hspace{1cm} (11)

where $\hat{y}_{ij}$ is the $j$th component of the discrete Fourier transform of the $i$th segment of data of length $M$, $y_i = \{y_k : k = 1 + (i-1)M, ..., iM \}$. $M$ is assumed to be even and to be a factor of $N$, so that $i = 1, ..., N/M$, and $j = 1, ..., M/2 + 1$. $j=1$ and $j= M/2 + 1$ correspond to DC and to the Nyquist frequencies, and will not be considered anywhere below. Consequently, the maximum number of useful pixels in the time-frequency representation will be $N_T = (M/2 - 1)(N/M)$. The time resolution $T$ of the time-frequency representation is fixed and is given by $T = M/f_s$. The frequency resolution $F$ is simply $F = 1/T$. In this time-frequency representation, the two hypotheses expressed in Eqs. (1) and (2) become $P_{ij}(y) = |\hat{s}_{ij} + \epsilon \hat{n}_{ij}|^2$ and $P_{ij}(y) = |\epsilon \hat{n}_{ij}|^2$, respectively.

If the noise $n$ is Gaussian and white, the power in different pixels is statistically independent. When no signal is present, $P_{ij}$ is the sum of the square of two independent Gaussian variables with zero mean and equal variance, and is therefore exponentially distributed:

$$p_{P_{ij}}(P) = \frac{e^{-P/\epsilon^2}}{\epsilon^2}. \hspace{1cm} (12)$$

When a signal is present, the Gaussian variables have non-zero mean, and the probability density function (hereafter, pdf) of the power is [22]:

$$p_{P_{ij}}(P|\hat{s}_{ij}|^2) = \frac{e^{-(P+|\hat{s}_{ij}|^2)/\epsilon^2}I_0 \left(2\sqrt{P|\hat{s}_{ij}|^2}/\epsilon^2 \right)}{\epsilon^2}, \hspace{1cm} (13)$$

where $I_0$ is the modified Bessel function of order zero.

The spectrogram representation is of course not the only time-frequency representation available, and may not be optimal for most signal. It is however the simplest one to work with in this exploratory work. Wavelet bases share the independence property of the noise in the different pixels, as do any orthonormal basis, and they are known to offer better localization properties than Fourier transforms for many signals [23]. However, their dyadic representation makes their analysis more complicated, due to the varying shape of the pixels with scale. Another classical way to improve the spectrogram representation is to use windows and to overlap the segments. This is a good way to reduce artifacts from Fourier transforms (e.g. edge effects) and to increase the time resolution of the time-frequency representation, but at the cost of destroying the statistical independence of the pixels.

Step (ii) of **TFCLUSTERS** consists in applying a threshold on the power $P_{ij}$. Pixels with $P_{ij} > \eta$ are called black pixels, and other pixels are called white pixels. The probability for any given pixel to be black when no signal is present (the black pixel probability, $p$), is given by:

$$p = \exp(-\eta/\epsilon^2). \hspace{1cm} (14)$$

Step (iii) of **TFCLUSTERS** considers the clustering of the black pixels. A cluster is defined as a set of black pixels containing all the black pixels that are the nearest neighbour of any pixel in the set. For a given pixel, its nearest neighbours are the pixels immediately to its left and to its right (time steps immediately before and after), and above and below it (frequency difference equal to the spectrogram resolution $F$). Two pixels touching only by a “corner” are called next nearest neighbours. The size $S$ of a cluster is simply defined as the number of black pixels it contains.
The notion of distance $d_c$ between two clusters $\Gamma_1$ and $\Gamma_2$ is defined as the minimal distance $d_p$ between any two pixels $p_1$ and $p_2$ in the two clusters,

$$d_p(p_1, p_2) = |i_1 - i_2| + |j_1 - j_2|,$$

(15)

$$d_c(\Gamma_1, \Gamma_2) = \min_{p_1 \in \Gamma_1, p_2 \in \Gamma_2} d_p(p_1, p_2),$$

(16)

where $(i_k, j_k)$ are the coordinates of $p_k$, i.e., $p_1$ corresponds to $P_{i_1j_1}(y)$, etc. Hence, nearest neighbour pixels have $d_p = 1$, next-nearest neighbour pixels have $d_p = 2$, and any two clusters must have $d_c \geq 2$. This choice of distance is made for convenience, but it has the implication of making the distance isotropic in the spectrogram representation of the time-frequency domain, irrespectively of the actual values of its time and frequency resolutions. Building a spectrogram with very long time bins and therefore very narrow frequency bins would have the effect of making the distances “longer” in the frequency direction than a spectrogram with short time bins and wide frequency bins.

Thresholds are applied both on the size of the cluster and on its distance to other clusters. The latter is easily motivated for physical signals; for example, although clusters of size two that are produced by fluctuations in the noise could be likely in a certain experiment, and therefore be below the size threshold, to have two or more such clusters close from each other, say being next nearest neighbours, could be rather unlikely. Hence, a signal with a well defined frequency that is slowly varying, so that its time-frequency track is a thin curve, could easily produce an archipelago of clusters of size two that is statistically unlikely to be produced by noise fluctuations alone.

If a cluster $\Gamma_1$ has size $S_1 \geq \sigma$, it is immediately accepted as significant. Otherwise, its distance to all other clusters $\Gamma_2$ with size $S_2 < \sigma$ is compared to a threshold $\delta_{S_1, S_2}$ which depends explicitly on the size of the two clusters. All the clusters with $d_c(\Gamma_1, \Gamma_2) \leq \delta_{S_1, S_2}$ are merged into a generalized cluster, which is declared significant. If $\Gamma_2$ is already in a separate generalized cluster, $\Gamma_1$ is added to that generalized cluster. If no cluster $\Gamma_2$ with $S_2 < \sigma$ satisfies the distance criterion, $\Gamma_1$ is rejected. For a given choice of the minimum cluster size $\sigma$, there are $\sigma(\sigma - 1)/2$ distance thresholds $\delta_{S_1, S_2}$, which will be organized below as a vector $\delta$:

$$\delta = [\delta_{S_1, S_2}] = [\delta_{1,1}, \delta_{1,2}, ..., \delta_{1,\sigma-1}, \delta_{2,\sigma-1}, ..., \delta_{\sigma-1,\sigma-1}].$$

(17)

Finally, step (iv) of TFCLUSTERS considers significant clusters from step (iii) and threshold on their excess power $\hat{Q}$, which is defined as

$$\hat{Q} = \sum_{(i,j) \in \Gamma} (P_{ij}(y) - \eta),$$

(18)

for any given cluster $\Gamma$ of size $S$. If no signal is present, i.e. if the cluster $\Gamma$ results from a fluctuation of the noise, the pdf of $P_{ij}(y)$ after the thresholding of step (ii) will be a truncated exponential for $(i, j) \in \Gamma$:

$$p_{P_{ij}}(y)(P) = \begin{cases} e^{-(P-\eta)/\epsilon_2} & \text{if } P > \eta \\ 0 & \text{otherwise}. \end{cases}$$

(19)

The pdf of $\hat{Q}$ will be the convolution of $S$ such distributions, i.e.

$$p_{\hat{Q}}(P) = \left\{ \begin{array}{ll} \frac{1}{(S-1)!} \left( \frac{P-S\eta}{\epsilon_2} \right)^{S-1} e^{-(P-S\eta)/\epsilon_2} & \text{if } P > S\eta \\ 0 & \text{otherwise.} \end{array} \right.$$

(20)

Hence, setting a size-dependent threshold $Q_S$ on $\hat{Q}$, defined by the integral equation

$$\alpha = \int_{S\eta}^{Q_S \eta} \frac{1}{(S-1)!} \left( \frac{P-S\eta}{\epsilon_2} \right)^{S-1} e^{-(P-S\eta)/\epsilon_2} dP,$$

(21)

leads when only noise is present to the rejection of a fraction $\alpha$ of the clusters that survive step (iii), independently of the cluster size $S$.

Step (iv) is very similar to the scheme used in the excess power algorithm of [15]. The essential difference is that TFCLUSTERS chooses the pixels included in the computation of the excess power by using a threshold on the individual pixel power and a clustering analysis, while [15] use a fixed set of “masks” of different shapes that are translated in the time-frequency plane, and over which the power for all pixels is integrated. It should also be noted that other kinds of thresholds on the total cluster power could be use in step (iv). While the one presented here penalizes in a similar fashion clusters of all sizes (it reduces the false rate by $\alpha$ independently of $S$), a size independent threshold could also be used, for instance, in which case large clusters would be more likely to survive the last step of TFCLUSTERS.
IV. OPERATING CHARACTERISTICS

Assessing the optimality of TFCLEUSTERS in the modified Neyman-Pearson sense is a very involved mathematical task, as it can be inferred from similar problems treated in the literature [24]. Nevertheless, it is shown in this section that the algorithm presents interesting properties that suggest its near-optimality for a large class of signals. What will actually be discussed is a simplified version of the algorithm not involving the clustering analysis (step (iii) of TFCLEUSTORS). It is assumed that the clustering analysis will improve the performances of this detector when signals that form clusters in the time-frequency plane are indeed present.

The binary test will be constructed by comparing the signal power estimate, $\hat{Q}(y) = |\hat{s}|^2$, to a certain threshold $\zeta$; $\hat{Q} > \zeta$ will lead to the acceptance of Eq. (1). This number is constructed by summing the power in the spectrogram over the pixels that have power above a certain threshold:

$$|\hat{s}_{ij}|^2 = \begin{cases} 0 & \text{if } P_{ij}(y) < \eta \\ P_{ij}(y) - \eta & \text{otherwise} \end{cases}$$  \hspace{1cm} (22)

$$\hat{Q}(y) = \sum_{i,j} |\hat{s}_{ij}|^2.$$  \hspace{1cm} (23)

The sum in Eq. (24) is performed over the whole time-frequency plane, and therefore is over $N/2$ terms.

The following theorem, inspired from the work of [21] on signal estimation, is proved in appendix A:

**Theorem 1:**

Given the model Eqs. (22)-(24), for $Q = |s|^2$, and provided that $\eta = \beta \epsilon^2 \log N/2$, $\beta > 1$, there exists a sequence of numbers $\pi_N$ with $\pi_N \to 1$ for $N \to \infty$ such that $\forall s \in \mathbb{R}^N$:

(i) $Pr(\hat{Q} \leq Q) = \pi_N$,

(ii) $Pr(\hat{Q} \geq \hat{q}) = \pi_N$, where $\hat{q}$ is any power estimator with $Pr(\hat{q} \leq Q') = \pi_N \forall s' \in \Omega(s)$, where $\Omega(s)$ is some neighborhood of $s$ defined in Eq. (A4), and $Q' = |s'|^2$.

(iii) $Pr\left(\hat{Q} \geq Q - \sum_{i,j} \min(2\beta \epsilon^2 \log N/2, |\hat{s}_{ij}|^2)\right) = \pi_N$.

This theorem leads to a number of observations:

- **Optimality:** Theorem 1 shows the optimality of $\hat{Q}$ in the subset of detectors for which the condition stated in part (i) is respected: $\hat{Q}$ provides the tightest lower bound to $Q$ that can be constructed from the data, for all $s$.

- **False alarm probability:** For a threshold $\zeta$ on $\hat{Q}$, part (i) implies that the probability to detect a signal with power $Q < \zeta$ goes to zero asymptotically with $N$ getting large. In particular, part (i) implies that the false alarm probability goes to zero asymptotically, since when no signal is present $Q = 0$.

- **False dismissal probability:** The direct interpretation of part (iii) is that any signal with $Q > \zeta + \sum_{i,j} \min(2\beta \epsilon^2 \log N/2, |\hat{s}_{ij}|^2)$, or equivalently with $\sum_{i,j} (|\hat{s}_{ij}|^2 - 2\beta \epsilon^2 \log N/2)_+ > \zeta$, will be detected with a false dismissal probability approaching zero. Obviously, if $\max |\hat{s}_{ij}|^2 \leq 2\beta \epsilon^2 \log N/2$, part (iii) implies that $\hat{Q} = 0$, giving a limit on how sensitive this detector can be.

- **Sparse signals:** Also as a consequence of part (iii), sparse signals that have only a small number of non-zero $\hat{s}_{ij}$ are more easily detected, because they have a smaller value of $\sum_{i,j} \min(2\beta \epsilon^2 \log N/2, |\hat{s}_{ij}|^2)$ than signals with the same power spread over a larger number of pixels.

The scheme discussed above of thresholding on the power $P_{ij}(y)$ integrated over pixels that have power above a certain threshold $\eta$ thus possesses certain optimal properties. It is very efficient for signals that have a sparse representation in the time-frequency domain. It is unlikely, however, that the optimal properties are preserved when the signal subspace is restricted to signals that form clusters in the time-frequency domain; it that case, intelligence about the spatial correlation of the signal pixels should certainly allow for more efficient algorithms. The idea behind TFCLEUSTORS is that the ad hoc approach of merging the two-threshold scheme discussed above with a clustering analysis designed to effectively reject noise should lead to an efficient algorithm for such a signal subspace.

V. CLUSTERING ANALYSIS

On physical grounds, it can be expected for most transient sources that the pixels with excess power due to the signal will tend to cluster in the time-frequency domain. Short signals like black hole ringdowns or mergers, for instance,
have durations roughly of the order of the inverse of the usable interferometer bandwidths (∼ 1 kHz), and therefore appear as connected clusters of duration T and bandwidth equal to the search bandwidth. For longer signals that spend tens or hundreds of cycles in the interferometers band, the required stability of the source is likely to necessitate that the dominating mechanism for the emission of the waves is governed by rotation, with the wave instantaneous frequency being around \(J/\pi f\), for \(J\) the magnitude of the angular momentum along the principal axis of rotation, and \(I\) the moment of inertia of the source. Complicated dynamics may lead to the spreading of the signal power over some finite frequency interval, or even the formation of sidebands well-separated from the principal frequency of the waves (in this case, each sideband will be considered an independent cluster). When the source is not too far away, it is argued below that the amount of angular momentum radiated by gravitational waves with amplitude near the sensitivity limit of the interferometric detectors is insufficient to cause a change in the rotation frequency that is rapid enough to produce disconnected pixels in the time-frequency representation of the signal.

Over a time interval \(T\) corresponding to one time slice in the time-frequency representation, a source at a distance \(r\) from the Earth that emits uniformly in all directions waves of frequency \(f\) and characteristic strain amplitude \(h\) will radiate a total amount of energy \(\Delta E\), where [2]

\[
\Delta E \sim \pi^2 c^3 G \frac{f^2}{3} h^2 r^2 T.
\] (25)

The change \(\Delta J\) in angular momentum magnitude corresponding to the emission of the waves is related to the amount of radiated energy by

\[
|\Delta J| \sim \frac{\Delta E}{\pi f}
\] (26)

if most of the radiation is quadrupolar [25]. On the other hand, for sources with dynamics dominated by rotation, the second time derivative of the mass quadrupole moment is bounded by \(\dot{Q} \leq \pi J f\), so for \(h = 2G\dot{Q}/r c^4\),

\[
J \geq \frac{c^4}{2\pi G} \frac{hr}{f}.
\] (27)

The value of the characteristic amplitude \(h\) is expressed in term of the noise spectral density \(S_n(f)\) so that the signal-to-noise ratio for a signal with bandwidth \(1/T\) is unity:

\[
h = S_n(f)/T^{1/2}.
\] (28)

Combining Eqs. (25)-(28):

\[
\frac{|\Delta J|}{J} \lesssim 3 \cdot 10^{-4} \left(\frac{r}{1 \text{ Mpc}}\right) \left(\frac{f}{100 \text{ Hz}}\right)^2 \left(\frac{S_n(f)}{5 \cdot 10^{-23} \text{ Hz}^{-1/2}}\right) \left(\frac{T}{0.1 \text{ s}}\right)^{1/2},
\] (29)

where the numbers in Eq. (29) for \(S_n(f)\) and \(f\) correspond approximately to the minimum of the noise spectral density of the interferometers being presently developed, and the value of \(T\) is chosen to match the expected time resolution of the time-frequency representations to be used on actual data. For values of \(r\) that are sufficiently small, the variation of the angular momentum magnitude over a time \(T \sim 0.1\) s is negligible. In order of magnitude, \(|\Delta J|/J \sim |\Delta f|/f\), for \(\Delta f\) the change in the wave frequency over a time \(T\), provided that the source doesn’t change its moment of inertia by a large fraction, which is unlikely to happen over timescales of 0.1 s, except perhaps near the end of the gravitational wave signal (as e.g. in binary inspirals, at the innermost stable circular orbit), at which point the question of the amount of clustering becomes irrelevant. Hence, the wave frequency is not expected to vary enough over the time resolution \(T\) to generate pixels in contiguous time slices that are disconnected. This can happen if

\[
\frac{|\Delta f|}{f} \gtrsim \frac{1}{T f}.
\] (30)

Combining Eqs. (29) and (30) gives a necessary (but not sufficient) condition for pixels to be disconnected:

\[
r \gtrsim 300 \text{ Mpc} \left(\frac{100 \text{ Hz}}{f}\right)^3 \left(\frac{5 \cdot 10^{-23} \text{ Hz}^{-1/2}}{S_n(f)}\right) \left(\frac{0.1 \text{ s}}{T}\right)^{3/2}.
\] (31)

In this context, any rotation dominated source at the detection limit of the detector that is closer than \(r\) is expected to form a cluster in the time-frequency domain. Figure 1 shows the variation of \(r\) with frequency, for the noise spectral
FIG. 1: The maximal distance $r$ at which rotation dominated sources at the sensitivity limit of the LIGO 4k interferometers are expected to form clusters in the time-frequency plane.

density of the LIGO interferometers [26]. Sources with $f \lesssim 300$ Hz will form clusters even if they are as far as the Virgo cluster; those with $f \gtrsim 1250$ Hz, even if they are in the galaxy, may not form clusters.

These considerations show that under certain restrictions, at least two broad classes of sources (short impulsive and rotation dominated) should lead to signals that form clusters in the time-frequency plane. As it will be shown below, Gaussian noise will have the opposite property, in the sense that black pixels will tend to fill the plane uniformly, without forming large clusters. This will be used as a powerful basis for denoising the thresholded spectrograms computed from the data.

A. False Alarm Rate

The reader familiar with the mathematics of percolation theory in statistical physics has most likely recognized at this point the applicability of the results in that field to the problem under investigation here. In particular, for a black pixel probability $p$, the average number of clusters of size $S$ per pixel of an infinite image is:

$$\langle n_S(p) \rangle = p^S D_S(1 - p),$$

where $D_S$ is the so-called perimeter polynomial, and is related to the number of shapes a cluster of size $S$ can have (counting shapes related only by a translation as identical), that is to the degeneracy $g_{SR}$ of a cluster of size $S$ and perimeter $R$:

$$D_S(q) = \sum_R g_{SR} q^R.$$  \hspace{1cm} (33)

As it can be seen from Eq. (33), the perimeter of a cluster is the number of white pixels having at least one black pixels in the cluster as a nearest neighbour. Coefficients of the perimeter polynomials for cluster sizes up to $S = 22$ for nearest neighbours on the square lattice are tabulated in [27], and were mostly generated through the use of computer enumeration techniques. Note that $\langle n_S(p) \rangle$ is simply the probability of any pixel to be in a cluster of size $S$, divided by $S$ [28].

For low cluster densities, the expected number of clusters per unit time $\lambda$, i.e. the cluster false alarm rate, is related for a threshold $\sigma$ on the cluster size to the time and frequency resolutions, and to the bandwidth $B$ that is searched:

$$\lambda = B \frac{1}{TT} \sum_{S=\sigma}^{\infty} \langle n_S(p) \rangle.$$  \hspace{1cm} (34)
FIG. 2: The mean number per pixel of clusters of size greater or equal to the threshold \( \sigma \) for different black pixel probabilities. Solid line: \( p = 0.1 \), dashed: \( p = 10^{-3} \), dash-dotted: \( p = 10^{-3} \), dotted: \( p = 10^{-4} \).

In practice, because of the rapid decay of \( \langle n_S(p) \rangle \) with \( S \), the sum can easily be truncated without major losses of precision. Figure 2 illustrates the nearly exponential decay of the cluster rate with the threshold on the cluster size, for different black pixel probabilities.

In analogy to Eq. (32), the average number per pixel of pairs of clusters of size \( S_1 \) and \( S_2 \) separated by a distance \( d \) is defined as:

\[
\langle \nu_{S_1,S_2}(d) \rangle = p^{S_1+S_2} H_{S_1,S_2}^d (1-p),
\]

(35)

where the polynomial \( H_{S_1,S_2}^d \) is related to the number of configurations \( k_{S_1,S_2,R}(d) \) for a cluster of size \( S_1 \) to be within a distance \( d \) from a cluster of size \( S_2 \), where the sum of the perimeters of the two clusters is \( R \), and where configurations that are related only by a translation are again considered identical:

\[
H_{S_1,S_2}^d(q) = \sum_R k_{S_1,S_2,R}(d) q^R.
\]

(36)

As an example, for two clusters of size 2 separated by a distance of two, Eq. (35) becomes:

\[
\langle \nu_{2,2}(2) \rangle = \frac{56p^4q^{11} + 40p^4q^{10}}{4}.
\]

(37)

The factors in the numerator account for all the possible configurations for a cluster of size 2 to be a distance 2 from another cluster of size 2, and the factor of 4 in the denominator is the total number of pixels, necessary in order not to overcount clusters related by a simple translation. The general expression for \( d > 2 \) is

\[
\langle \nu_{2,2}(d) \rangle = 8(d+1)p^4q^{12},
\]

(38)

and hence the associated false alarm rate is

\[
\lambda_{2,2} = \frac{B}{TF} \left[ 10p^4q^{10} + 14p^4q^{11} + \sum_{d=3}^{\delta_{2,2}} 8(d+1)p^4q^{12} \right]
\]

(39)

\[
= \frac{B}{TF} \left[ 10p^4q^{10} + 14p^4q^{11} + 4(\delta_{2,2}^2 + 3\delta_{2,2} - 10)p^4q^{12} \right], \text{ if } \delta_{2,2} \geq 3,
\]

(40)

where \( \delta_{2,2} \) is the threshold on the distance for two clusters of size 2 to be considered correlated.
<table>
<thead>
<tr>
<th>$S_1$</th>
<th>$S_2$</th>
<th>$d$</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>24</td>
<td>36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td>80</td>
<td>128</td>
<td>64</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>240</td>
<td>128</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>304</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>368</td>
<td>192</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>432</td>
<td>224</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td>260</td>
<td>480</td>
<td>360</td>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>720</td>
<td>680</td>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>900</td>
<td>840</td>
<td>220</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>1080</td>
<td>1000</td>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>1260</td>
<td>1160</td>
<td>300</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td>56</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>128</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td>240</td>
<td>400</td>
<td>160</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>680</td>
<td>360</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>840</td>
<td>440</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>1000</td>
<td>520</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>1160</td>
<td>600</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td>792</td>
<td>1344</td>
<td>984</td>
<td>216</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>1944</td>
<td>1824</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>2376</td>
<td>2208</td>
<td>576</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>2808</td>
<td>2592</td>
<td>672</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>3240</td>
<td>2976</td>
<td>768</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td>12</td>
<td>324</td>
<td>684</td>
<td>456</td>
<td>108</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>864</td>
<td>912</td>
<td>240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>1056</td>
<td>1104</td>
<td>288</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>1248</td>
<td>1296</td>
<td>336</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>1440</td>
<td>1488</td>
<td>384</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td>112</td>
<td>2016</td>
<td>4424</td>
<td>4172</td>
<td>1736</td>
<td>280</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>4788</td>
<td>7000</td>
<td>3528</td>
<td>616</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>5796</td>
<td>8400</td>
<td>4200</td>
<td>728</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>6804</td>
<td>9800</td>
<td>4872</td>
<td>840</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>7812</td>
<td>11200</td>
<td>5544</td>
<td>952</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td></td>
<td></td>
<td>272</td>
<td>2912</td>
<td>7216</td>
<td>8112</td>
<td>4800</td>
<td>1440</td>
<td>176</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td>6480</td>
<td>12096</td>
<td>8800</td>
<td>2944</td>
<td>384</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>7776</td>
<td>14400</td>
<td>10400</td>
<td>3456</td>
<td>448</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td>9072</td>
<td>16704</td>
<td>12000</td>
<td>3968</td>
<td>512</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td>10368</td>
<td>19008</td>
<td>13600</td>
<td>4480</td>
<td>576</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE I: Coefficients $k_{S_1S_2R}(d)$, multiplied by $S_1 + S_2$, for clusters of size $S_1$ and $S_2$, with total perimeter $R$, and separated by a distance $d$.  

131
Table II: General formulae for $\langle \psi_{S_1,S_2}(d) \rangle$ for distances $d > 2$, for different cluster sizes $S_1$ and $S_2$, as deduced from Table I. By definition, $q \equiv 1 - p$.

Table I gives the coefficients $k_{S_1,S_2}(d)$ for small cluster sizes that were obtained by computer enumeration. General relations can easily be deduced from this table: on the square lattice, using the definition of distance from Eq. (15), the diameter of a circle of radius $r$ is $4r$, and therefore the number of points at distance $d$ from a certain cluster increases linearly with $d$. Moreover, when $d > 2$, both clusters are guaranteed to not be sharing any perimeter white pixels, and therefore the number of configurations with fixed total perimeter consists of a “bulk” part (from the constant number of configurations occurring at pixels that are on the same row or column as a pixel in the cluster), and of a part growing linearly with $d$ (from the “diagonals” of the curve of constant distance from the first cluster). Table II gives the formula for $\langle \psi_{S_1,S_2}(d) \rangle$ deduced from Table I, for $d > 2$. Values for $d = 2$ can be read directly from Table I. Figure 3 shows the increase of the rate of correlated clusters with the threshold on the distance between the two clusters.

It should be noted that TFCLUSTERS includes higher order terms ($n$-cluster configurations, $n \geq 3$) into its definition of generalized clusters. These complex configurations are built by merging all the clusters satisfying the distance criteria, and consequently the sum of Eq. (34) and of its equivalent for Eq. (35) overestimates slightly the true false alarm rate, especially when the distance thresholds $\delta_{S_1,S_2}$ are large for the density of small clusters.

B. Efficiency

As shown above, the false alarm rate is computed analytically; when $\alpha = 0$ (cf. step (iv) of TFCLUSTERS), the probability of detection of a certain signal can also be computed analytically. It is not possible at this time to modelize the detection efficiency when $\alpha > 0$ because the equivalent to Eq. (20) when a non-null signal is present is not known analytically, and is required to compute the probability of detection of a given signal. The sensitivity and false alarm rate of TFCLUSTERS depend on a number of parameters: $F, T, p, \sigma$ and $\delta$. Using some general information about the expected waveform, it is possible to run an optimization over all these parameters to maximize the probability of detection at fixed false alarm rate. More generally, it is possible to run multiple versions of the detector, each one for a different set of parameters, in order to cover as many different classes of signal as possible. The number of classes that can be considered is limited by the computational power available for the analysis, and is somewhat limited by the fact that results are not independent.

For comparison, the probability of detection using the ideal power detector can also be computed. This detector assumes that the bandwidth and duration of the signal are known, and it compares the excess power computed according to these parameters to a threshold $\kappa$. In order to get an false alarm rate rather than a false alarm probability, it is assumed that the detector is applied with frequency $1/\tau$ on segments of independent data of length $\tau$, and that all frequency bins in the search bandwidth $B$ are considered in disjoint groups having bandwidth $\phi$. This detector is incoherent and is therefore not as efficient as match filtering, but it still performs better than any detector that can be implemented when neither the duration nor the bandwidth is known. The false alarm rate is:

$$
\lambda = \frac{B}{\phi \tau} \int_{\kappa}^{\infty} \frac{(P/e)^{u-1} e^{-P/e^2}}{e^{2(uw-1)}} \, dP
$$

(41)
and the probability of detection is [22]

\[
\beta = \int_{\kappa}^{\infty} \left( \frac{P}{uvp^2\epsilon^2} \right)^{(uv-1)/2} \exp \left( -\frac{P + uv\rho^2\epsilon^2}{\epsilon^2} \right) I_{u-1} \left( 2\sqrt{\frac{uv\rho^2P}{\epsilon^2}} \right) \frac{dP}{\epsilon^2}.
\]

(42)

The case where one or more of the bandwidth, duration, central frequency or arrival time parameters have to be searched over is more complicated to analyze: the probabilities are not statistically independent, because of the need to search over overlapping regions in the time-frequency representation of the signal, and therefore the computation of the false alarm rate is rather involved.

An example is now presented to make things more precise regarding how well TFCLUSTERS compares to the ideal power detector. The signal is taken to be of duration 6T and bandwidth smaller than \( F = 1/T \) (since the problem is symmetric under the interchange of the time and frequency axes, the results below are equivalent to the case of a short signal of duration \( T \) and bandwidth \( 6/T \)). It is assumed that the signal power is distributed uniformly over its full duration and bandwidth. For definitiveness the bandwidth of the search is chosen to be \( B = 512F \), and the false alarm rate is \( \lambda = 1/3600T \). The central frequency is assumed to be such that the power is concentrated in a single constant frequency row of the spectrogram, while both the case of synchronized (the signal covers six pixels) and the case of random arrival time (the signal covers seven pixels) are considered. For the ideal power detector, \( uv = 6 \) or 7, and \( \tau \) is set to 6T. For the TFCLUSTERS algorithm, the various thresholds (\( \eta, \sigma \) and \( \delta_{s_1}, s_2, \alpha \) being zero) are optimized for every value of the signal-to-noise ratio in order to maximize the probability of detection (hereafter, POD) for the
constraint $\lambda \leq 1/3600T$. The details of the calculation are presented in appendix B. Figure 4 shows a comparison of the optimized POD as a function of signal to noise ratio $\rho$, for TFCLUSTERS and the ideal power detector. Only two sets of clustering analysis thresholds cover the whole range of signal-to-noise ratios of figure 4, illustrating the relative independence of the performances of TFCLUSTERS on its numerous parameters. Given the fact that the chosen signal was a line in the time-frequency plane, and was therefore the shape that is the easiest to “break” by changing a black pixel into a white pixel, the example presented here can be considered as a “worse case” situation; other distributions of the power (i.e. other signal shapes) would make the performances of TFCLUSTERS and of the ideal power detectors only closer.

VI. NUMERICAL SIMULATIONS

Extensive numerical simulations were carried out in order to confirm the validity of the analyses presented in the previous sections, and in order to explore properties of TFCLUSTERS that are hard to study analytically. All of the results below were produced using the same method: segments of Gaussian white noise of unit variance are produced using a random number generator [29]. The segments are of duration $10^4$ s, and are sampled at 2048 Hz. These numbers are chosen so that processing a single segment uses most of the RAM of the computers on which the code is running. On every segment, the implementation of TFCLUSTERS within the LIGO Scientific Collaboration Algorithm Library [30] is used to generate a list of significant clusters, according to some pre-specified values of the various thresholds. The 90% central confidence interval for the rate of significant clusters in the data, assuming they form a Poisson process, is computed based on the number of detected clusters, using the standard Neyman construction [31]. If the ratio of the width of the 90% confidence interval to its central value is smaller than 1%, the simulation terminates. Otherwise, a number of other $10^4$ s segments are processed, until the above termination criterion is met. For a true rate of $\lambda$, the termination criterion requires to process of order $10^9/\lambda$ worth of simulated data. It should be noted that the timing distribution of clusters from white noise is indeed very well approximated by a Poisson distribution. This is confirmed in one specific case by figure 5 which shows that the distribution of the time delay $\Delta$ between two successive clusters follows an exponential distribution, which is the defining property of a Poisson process.

The simulations were performed on a cluster of workstations with 1GHz Intel Pentium III “Coppermine” processors, 512 Mbytes of PC-133 RAM memory, with the Linux operating system. For a large range of parameter values, the
FIG. 5: The probability density \( p(\Delta) \) of the time delay \( \Delta \) between two successive clusters, as measured empirically for \( \eta = 3.719 e^2 \), \( \sigma = 5 \), \( \delta = [0, 0, 0, 0, 0, 0, 2, 3, 4] \), and \( \alpha = 0 \), for \( T = 1/32 \) s, \( B = 992 \) Hz. The two continuous lines correspond to the extrema of the predicted Poisson distribution, assuming the value of the rate at both ends of its 90\% confidence interval \((7.28 \pm 0.02) \cdot 10^{-3} \) Hz.

processing time per CPU was on average 250 to 550 times shorter than the duration of the data segment. Most of the time was spent at grouping the black pixels into clusters, and quite logically the most important factor in determining the speed of tfclusters was the pixel black probability; independently of the other parameters, the ratio of processing to real time was around 300 for \( \eta/e^2 = 2 \), around 500 for \( \eta/e^2 = 4 \), and was increasing almost linearly with the power threshold \( \eta \).

A. The Case \( \delta = 0, \alpha = 0 \)

When \( \delta = 0 \) and \( \alpha = 0 \), the expected rate is given by Eq. (34). Figures 6 and 7 show the excellent agreement between rates from simulations and predictions from Eq. (34). The two agree to better than 0.5\% most of the time, commensurably with the precision of the simulations. The sum in Eq. (34) is of course dominated by the first few terms, but values of \( \langle n_S(p) \rangle \) as predicted from Eq. (32) describe the simulated data also very well for large values of \( S \), as shown in figure 8.

B. The Case \( \delta \neq 0, \alpha = 0 \)

For \( \delta \neq 0 \), there is a contribution to the cluster rate from both Eq. (34) and its equivalent for \( \langle \nu_{S, S_2}(d) \rangle \). Figures 9 and 10 show the good agreement between rates from simulations and predictions. When \( \eta/e^2 \lesssim 2.8 \), the agreement is at the precision level of the simulations, although there is a systematic overestimation of the measured rate by the predictions for \( \eta/e^2 \lesssim 3.7 \). This overestimation, however, reaches almost 20\% for \( \eta/e^2 = 2 \); this is expected, as explained in section V, because for \( \eta/e^2 \lesssim 2.8 \) the cluster density is high enough that higher-order combinations above the 2-cluster one are likely to be produced.

To demonstrate that the error is indeed due to higher-order terms, a histogram of the contribution \( \lambda(S) \) to the total rate of generalized clusters of size \( S \) is presented in figure 11, together with the theoretical prediction based on 1- and 2-cluster configurations. The predictions systematically overestimate the measured rate for \( S < 8 \), and underestimate it for \( S \leq 8 \). It should be noted that for \( \delta = [0, 0, 0, 0, 0, 0, 2, 3, 4] \), 3-cluster configurations can have sizes \( 8 \leq S \leq 12 \), while 4-cluster configurations can have sizes \( 10 \leq S \leq 16 \). Hence, one can expect a small error for \( S = 8 \) corresponding to the 3-cluster \((S_1, S_2, S_3) = (2, 2, 4)\), and no errors from higher-order terms for \( S < 8 \), i.e. for the terms dominating the sum leading to the prediction for \( \lambda \). Overall, more small clusters than expected get merged into generalized clusters due to these high-order terms, and the total rate is consequently overestimated. Figure 12
FIG. 6: The measured rate $\lambda$ in white noise for various values of $\eta$, for $\delta = 0$, $\alpha = 0$, $T = 1$ s, $B = 1023$ Hz. From top to bottom, the curves correspond to $\sigma = 3, 4, 5, 7$. Both the error bars on $\lambda$ and the predicted rates from Eq. (34) are occulted by the thickness of the line.

FIG. 7: The fractional residuals $\Delta \lambda = (\lambda_{\text{pred}} - \lambda)/\lambda_{\text{pred}}$, where $\lambda_{\text{pred}}$ are the predicted rates, corresponding to figure 6. X-marks, plus signs, circles and squares correspond respectively to $\sigma = 3, 4, 5, 7$.

quantifies the importance of high-order terms as a function of the cluster density, i.e. as a function of the threshold $\eta$; as expected, the relative importance of high-order terms becomes smaller than the 1% level around $\eta/\epsilon^2 = 3$, in agreement with the results in figure 9.

C. The Case $\alpha > 0$, and Finite-Size Effects

For $\alpha > 0$, the rate is expected to be $(1 - \alpha)\lambda_0$, where $\lambda_0$ is the value of the rate for the same parameters, except that $\alpha = 0$. Figure 13 illustrates the reduction in $\lambda$ as $\alpha$ is varied. The results are as expected from Eq. (21), within
FIG. 8: The measured rate $\lambda(S)$ as a function of the cluster size $S$ for $\eta = 3e^2$, $\sigma = 5$, $\delta = 0$, $\alpha = 0$, $T = 1$ s, $B = 1023$ Hz. The plus signs are results from simulations, and the continuous line corresponds to the predictions. This plot was constructed from $7 \cdot 10^6$ s worth of simulated data.

FIG. 9: The measured rate $\lambda$ in white noise for various values of $\eta$, for $\sigma = 5$, $\delta = [0, 0, 0, 0, 0, 2, 3, 4, 4]$, $\alpha = 0$, $T = 1$ s, $B = 1023$ Hz. The dotted line correspond to the predicted rate.

the errors from the simulations.

The simulations presented in figure 13 were carried out with a very asymmetrical spectrogram: $T$ was chosen to be $1/32$ s, so $F = 32$Hz, and the 992 Hz bandwidth was covered by only 31 pixels, while a segment had $32 \cdot 10^4$ bins in time. Because of that, significant finite size effects are expected; the rate predictions are based on the assumption that the time-frequency plane is infinite, but in the present case it is small enough in the frequency dimension that clusters are likely to be "clipped" and therefore to be smaller then expected. Consequently, it is expected that the predictions will again overestimate the measured rate, and this is what is observed in figures 13 and 14. As it can be seen from figure 14, the fractional error from the prediction is still smaller than about 10% when the cluster density is low enough.
FIG. 10: The fractional residuals defined as in figure 7, but corresponding to figure 9.

FIG. 11: The measured rate $\lambda(S)$ as a function of the cluster size $S$ for $\eta = 3\sigma^2$, $\sigma = 5$, $\delta = [0, 0, 0, 0, 0, 0, 3, 4, 4]$, $\alpha = 0$, $T = 1$ s, $B = 1023$ Hz. The plus signs are results from simulations, and the continuous line corresponds to the predictions. This plot was constructed from $5 \times 10^9$ s worth of simulated data.

VII. CONCLUSIONS

**TFCLUSTERS**, a new time-frequency detector for bursts of gravitational radiation in broad-band interferometric observatories, was described in some details in this paper. The behavior of the detector when applied to white Gaussian noise in the absence of signals was carefully analyzed, leading to a formalism for the computation of the false alarm rate of **TFCLUSTERS** for any values of its many parameters. The results from numerical simulations showed that this analysis was accurate in most situations: errors at the $1\%$ level or better were obtained in “ideal” situations (low cluster density, large number of frequency bins), and errors of the order of $10\%$ appeared in situations where the analysis was expected to be less accurate, due to high order terms not included in the sums over the cluster configurations, or due to finite size effects from the limited bandwidth of the search. In the case where the errors
FIG. 12: The ratio $f$ of the number of generalized clusters containing three or more simple clusters to the total number of clusters, as a function of the threshold $\eta$, for the same parameters as in figure 9.

FIG. 13: The measured rate $\lambda$ in white noise for various values of $\eta$, for $\sigma = 5$, $\delta = [0, 0, 0, 0, 0, 0, 2, 3, 4, 4]$, $T = 1/32 \text{ s}$, $B = 992 \text{ Hz}$. From top to bottom, the curves correspond to $\alpha = 0, 3/4, 15/16$. The dotted line correspond to the predicted rate.

were large, the analysis presented in this paper was systematically overestimating the false alarm rates; should they prove from practical work to be necessary, more accurate estimates could therefore be obtained from more careful calculations, or from larger scale numerical simulations. While the false alarm properties of TFCLUSTERS are well-understood, the efficiency of the detector is subject to more uncertainties.

A calculation that was presented in Appendix B showed that the efficiency of TFCLUSTERS was comparable to that of the ideal power detector, unfortunately illustrating at the same time the mathematical complexity associated with producing such an estimate of the efficiency of the detector for a given signal. Nevertheless, the fundamental approach used by TFCLUSTERS, namely the use of an adaptive power integral over pixels with excessive power, was shown to be maximizing asymptotically the estimate of the power in the signal, provided it is not overestimating it, over all possible estimators of the same quantity. This naturally suggests optimal properties for TFCLUSTERS as a detector.
FIG. 14: The fractional residuals defined as in figure 7, but corresponding to figure 13, for the curve with $\alpha = 0$.

However, this can be nothing more than a conjecture at this point, as the actual proof of the optimality of TFCLUSTERS in the (modified) Neyman-Pearson sense is of great difficulty.

Independently of the question of the optimality of TFCLUSTERS, the structure of the algorithm is particularly practical for its implementation for the analysis of actual data. The first power threshold on individual pixels (step (ii)) can be chosen to be frequency dependent in order to allow the analysis of colored Gaussian noise when whitening of the data is not convenient; the errors on the black pixel probability that are introduced by the non-zero correlations coloring the noise are generally negligible. Moreover, frequency bands containing spurious interferences are easily left out of the analysis, with minor modifications to the algorithm presented in this paper.

The search for gravitational waves will of course require the operation of TFCLUSTERS in coincidence at sites that are geographically separated. While the comparison of event lists from individual sites is a possible way to carry this coincidence analysis, it may not give the best efficiency, in part because of the relatively coarse time resolution of TFCLUSTERS. Enhanced versions of the algorithm, operating on cross-spectrograms built from data from two interferometers, for instance, might be more efficient. The question of the integration of TFCLUSTERS into an analysis system designed to reject spurious interferences is currently the subject of active research[32].

Acknowledgments

The author wishes to thank Rainer Weiss for helpful discussions and comments on the work presented here, and the LSC group at the University of Wisconsin-Milwaukee for generous time allocation on their computer system. This work was supported by the National Science Foundation under cooperative agreement PHY-9210038, and by the Fonds pour la Formation de Chercheurs et l'Aide à la Recherche of the Province of Québec, Canada.

APPENDIX A: PROOF OF THEOREM 1

Consider first the deterministic equivalent to Eq. (1):

$$y = s + \delta u,$$  \hspace{1cm} (A1)

where $u$ is a nuisance parameter, so that the equivalent to Eq. (11) is

$$P_{ij}(y) = |\delta_{ij} + \delta \bar{u}_{ij}|^2,$$  \hspace{1cm} (A2)

where $\bar{u}_{ij} \in C$ and where by definition $|\bar{u}_{ij}| \leq 1$. 

140
Lemma 1:
For the model described by Eq. (A1) with \( \delta^2 = \eta, \forall s \in \mathbb{R}^N \) and \( \forall y \) satisfying Eq. (A1), \( \dot{Q}(y) \leq Q \).

Proof:
Consider Eq. (22). Trivially, if \( |\hat{s}_{ij}|^2 = 0 \), then \( |\hat{s}_{ij}|^2 \leq |\hat{s}_{ij}|^2 \). For \( |\hat{s}_{ij}|^2 > 0 \), \( |\hat{s}_{ij}|^2 = P_{ij}(y) - \eta \). From Eq. (A2), \( |\hat{s}_{ij}|^2 \geq P_{ij}(y) - \delta^2 \). Since \( \delta^2 = \eta \), this gives \( |\hat{s}_{ij}|^2 \geq |\hat{s}_{ij}|^2 \) for all \( s, y \). Summing over \( i \) and \( j \) gives \( \dot{Q} \leq Q \).

Lemma 2:
Given Eq. (A1) with \( \delta^2 = \eta, \forall s \in \mathbb{R}^N \), and \( \forall y \) respecting Eq. (A1), \( \dot{Q}(y) \geq \dot{q}(y) \), where \( \dot{q}(y) \) is any power estimator satisfying

\[
\dot{q}(y') \leq Q, \forall s' \in \Omega(s) \text{ and } \forall y' \text{ respecting Eq. (A1)},
\]

where

\[
\Omega(s) = \left\{ s' \in \mathbb{R}^N : |s - s'|^2 \leq \min\left(4\delta^2, 2|\tilde{s}_{ij}| \left( 1 + \frac{|\tilde{s}_{ij}|}{\delta} \left( 1 + \sqrt{1 + \frac{2\delta}{|\tilde{s}_{ij}|}} \right) \right) \right\}.
\]

Proof:
Suppose \( \exists y^0 \) such that \( \dot{q}(y^0) > \dot{Q}(y^0) \), so \( \tilde{y}_{ij}^0 = \tilde{s}_{ij}^0 + \delta \tilde{u}_{ij}^0 \). One can construct a signal \( s' \) such that

\[
\tilde{s}_{ij}' = \begin{cases} 
\tilde{y}_{ij}^0 - \delta \tilde{u}_{ij}^0 & \text{if } |\tilde{y}_{ij}^0| > \delta \\
0 & \text{otherwise.}
\end{cases}
\]

The freedom provided by \( \tilde{u}_{ij}' \in \mathbb{C} \) is used to choose all values of \( \tilde{u}_{ij}' \) so that \( \tilde{u}_{ij}' \) and \( \tilde{s}_{ij}' \) are orthogonal. If \( \tilde{s}_{ij}' = 0 \), the direction of \( \tilde{u}_{ij}' \) is unimportant. For all \( i, j, |\tilde{u}_{ij}'| = 1 \). Note that this choice of \( \tilde{u}_{ij}' \) and \( \tilde{s}_{ij}' \) satisfies Eq. (A1).

Also, since \( \tilde{s}_{ij}' - \tilde{y}_{ij}^0 = \delta (\tilde{u}_{ij}' - \tilde{u}_{ij}^0) \), \( |\tilde{s}_{ij}' - \tilde{y}_{ij}^0|^2 \leq 4\delta^2 \). Now, \( |\tilde{y}_{ij}^0|^2 = \delta^2 + |\tilde{s}_{ij}'|^2 \) so,

\[
|\tilde{s}_{ij}' - \tilde{y}_{ij}^0|^2 = |\tilde{s}_{ij}'|^2 + |\tilde{y}_{ij}^0|^2 - 2|\tilde{s}_{ij}'||\tilde{y}_{ij}^0| \leq 2|\tilde{s}_{ij}'|^2 + 2|\tilde{y}_{ij}^0| \delta + 2|\tilde{s}_{ij}'|^2 \left( 1 + \frac{2\delta}{|\tilde{s}_{ij}'|^2} \right).
\]

Combining these inequalities and summing over \( i, j \) shows that \( s' \in \Omega(s) \).

Then,

\[
\dot{Q}(y^0) = \sum_{i,j} (P_{ij}(y^0) - \eta) + \sum_{i,j} (|\tilde{s}_{ij}' + \delta \tilde{u}_{ij}'|^2 - \eta) + \sum_{i,j} (|\tilde{s}_{ij}'|^2 + \delta^2 - \eta)
\]

\[
= \sum_{i,j} (|\tilde{s}_{ij}'|^2) - |s'|^2.
\]

It follows that \( \dot{q}(y^0) > \dot{Q}(y^0) \) implies \( \dot{q}(y^0) > |s'|^2 \), contradicting Eq. (A3).

Lemma 3:
For the model described by Eq. (A1) with \( \delta^2 = \eta, \forall s \in \mathbb{R}^N \) and \( \forall y \) satisfying Eq. (A1), \( \dot{Q}(y) \geq Q - \sum_{i,j} \min(2\delta^2, |\tilde{s}_{ij}|^2) \).

Proof:
Obviously, \( |\tilde{s}_{ij}|^2 \leq |\tilde{s}_{ij}|^2 \). Also, from Eq. (A2), \( P_{ij}(y) \geq |\tilde{s}_{ij}|^2 - \delta^2 \), so from Eq. (22), \( |\tilde{s}_{ij}|^2 = P_{ij}(y) - \delta^2 \geq |\tilde{s}_{ij}|^2 - 2\delta^2 \) when \( |\tilde{s}_{ij}|^2 > 0 \). When \( |\tilde{s}_{ij}|^2 = 0, \delta^2 > P_{ij}(y) \geq |\tilde{s}_{ij}|^2 - \delta^2 \), so \( |\tilde{s}_{ij}|^2 < 2\delta^2 \). Combining the three inequalities gives \( |\tilde{s}_{ij}|^2 - |\tilde{s}_{ij}|^2 \leq \min(2\delta^2, |\tilde{s}_{ij}|^2) \); summing over \( i, j \) proves the lemma.

Proof of theorem 1:
In the noise model from Eq. (11), the power in the noise is distributed exponentially:

\[
p_{\delta^2}(P) = \frac{e^{-P/\delta^2}}{e^2}.
\]

141
The corresponding cumulative probability distribution is

\[ P_{\tilde{\eta}_{ij}}(P) = 1 - e^{-P/\epsilon^2}. \] (A13)

Hence, the probability that the maximum of \( N \) independent realizations of that random variable is less than a threshold \( C_N \) is

\[ \pi_N = P\left( \max_{i,j} |\tilde{\eta}_{ij}|^2 < C_N \right) = \left( 1 - e^{-C_N/\epsilon^2} \right)^N. \] (A14)

For \( N \gg 1 \), this can be rewritten as

\[ \pi_N = \exp \left( -e^{\log N - C_N/\epsilon^2} \right) \] (A15)

for \( C_N/\epsilon^2 > \log N \). Hence, for \( C_N \rightarrow \epsilon^2 \log N \), this probability goes to \( 1/e \) as \( N \) becomes large. Moreover, any values of \( \pi_N > 1/e \) can be achieved with the right choice of \( C_N \); for \( C_N = \beta \epsilon^2 \log N \), \( \beta > 1 \), \( \pi_N \rightarrow 1 \) for \( N \gg 1 \).

It should be noted that the event

\[ \max_{i,j} |\tilde{\eta}_{ij}|^2 < \beta \epsilon^2 \log N \] (A16)

implies a realization of the noise which can be mimicked by the deterministic noise model [Eq. (A1)] with \( \delta = \epsilon \sqrt{\beta} \log N \). Hence, the three statements of theorem 1 follow from the three lemmata proved in this appendix.

**APPENDIX B: THE NARROW-BAND SIGNAL EXAMPLE**

Given a \( u \) by \( v \) matrix \( Q \) with elements representing the distribution of the power in the signal in a rectangular sub-region of the time-frequency plane, the signal-to-noise ratio is defined by:

\[ \rho^2 = \frac{\sum_{i,j} Q_{ij}}{uv \epsilon^2}. \] (B1)

The elements of the matrix \( D \) representing the black pixel probability corresponding to the matrix \( Q \) are given by the integral of the density from Eq. (13):

\[ D_{ij} = \int_0^\infty p(P|Q_{ij})dP. \] (B2)

In general, not all pixels where some signal power is present will be black, and the signal will be detected only when a number of black pixels equal or larger than the threshold \( \sigma \) form a connected cluster, or when a pair of smaller clusters are close enough. The contribution to the probability of detection of the signal of such a configuration will be the product of the black pixel probabilities \( D_{ij} \) and of the white pixel probabilities \( 1 - D_{ij} \) for “holes”. Although noise fluctuations could in principle help the detection of a signal by forming a “bridge” over regions where no signal is present, summing over these possibilities involve \( 2^{uv} \) terms. A slight underestimate of the probability of detection is instead used by summing only over the \( n \) pixels where signal is present \( (n \leq uv) \), which reduces the number of terms to be considered to

\[ \sum_{n_H=0}^{n-\sigma} \frac{n!}{(n-n_H)!n_H!}, \] (B3)

where \( n_H \) is the number of holes. Of course, the enumeration process can be greatly simplified when the distribution of the power in the signal has some specific symmetries.

Consider now the example of section V B. Under the assumption that the starting time of the signal matches exactly the binning of the spectrogram used to detect it, it will be represented by a row of 6 pixels, each with an equal probability \( p \) to be black, neglecting effects such as power leakage. The columns labeled “POD \( (r \rightarrow p, s \rightarrow 0) \)” in tables III and IV give the contributions to the probability of detection (POD) of this signal from various thresholds. The problem is slightly more complex when the arrival time is taken to be random. In that case, the signal is spread
in general over 7 pixels, with the central five having a black pixel probability \( p \), and the leftmost and rightmost having smaller probabilities \( r \) and \( s \), respectively. \( r \) and \( s \) are given by:

\[
r = \int_{-\infty}^{\infty} p(P|P_s)dP
\]

and

\[
s = \int_{-\infty}^{\infty} p(P|P_s - P'_s)dP,
\]

where \( P_s \) is the power in the central five pixels, \( P'_s \) is the power in the leftmost pixel, and \( p(P|Q) \) is given by Eq. (13). The POD is then given by

\[
\frac{1}{P_s} \int_{0}^{P_s} POD(r(P'_s), s(P_s - P'_s))dP'_s,
\]

where POD\((r(P'_s), s(P_s - P'_s))\) is the probability of detection for the 7 pixels configuration; contributions to it from various thresholds can be found in the columns labeled “POD” in tables III and IV. For the ideal power detector with \( \lambda = 1/3600T \), \( \tau = 6T \) and \( B/\phi = 512 \), Eq. (41) gives \( \kappa \approx 23.87e^2 \) for \( uw = 6 \) and \( \kappa \approx 25.55e^2 \) for \( uw = 7 \). Eq. (42) is then used directly to generate the ideal power detector curves of figure 4.

---


[10] An arbitrary random process can always be reduced to a set of uncorrelated coefficients ("whitened") by the Karhunen-Loève expansion. Signals are affected by the whitening; however, the only place where this could have an adverse effect here is if the whitening were reducing the amount of clustering of astrophysical signals.
$S_1 \quad S_2 \quad \delta_{S_1,S_2} \quad \text{POD (} r \rightarrow p,s \rightarrow 0\text{)} \quad \text{POD}$

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>3</td>
<td>2</td>
<td>$p^3qrst$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>2</td>
<td>$2p^4qrst$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>2</td>
<td>$2p^4q(r + s - 2rs)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>$2p^3q^4rs$</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>2</td>
<td>$2p^4q^2(2p(1 - r)(1 - s) + q(r + s))$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>2</td>
<td>$2p^4q^2(r + s - 2rs)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>$2p^3q^2rs$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>2</td>
<td>$2p^4qrst$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>$2p^4qrst(r - 2s) + 2s)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3</td>
<td>$2p^3qrst$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>$3p^3q^2(1 + q) + p^2q^2(2r(1 - s) + 3e + p(4 - r - s - 2rs))$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>2</td>
<td>$2p^3qrst(r + s - 2rs)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>$2p^3qrst$</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>$4p^2qrst$ $pq^2(2p(1 - r)(1 - s) + q(r + s) + pq(3 - rs))$</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>2</td>
<td>$2p^2qrst(r + s - 2rs)$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>5</td>
<td>$2p^2qrst$</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td>$2pqrst$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>$q^3rs$</td>
<td></td>
</tr>
</tbody>
</table>

**TABLE IV:** Same as Table III for thresholds on the distance between two clusters of size $S_1$ and $S_2$.

[11] The Neyman-Pearson lemma shows that Eq. (3) is optimal when choosing between Eqs. (1) and (2); for the model $y(t) = An(t) + cn(t)$ with $s(t) = 1$, Eq. (3) is the most efficient estimator for $\lambda$, in the sense of being unbiased and of achieving the minimum possible mean square error (from the Fisher information, this is $\varepsilon^2$): $\hat{\lambda} = A - \varepsilon^2\partial \log p[y(t) | A] / \partial A - (y(t), s(t))$.


[20] Rules to use the Itô integral $\int_{0}^{1} b(t)u(t)dt = \int_{0}^{1} b(t)du(t)$ are similar to those for normal integration, except that second order terms in the increment $du(t)$ are preserved, corresponding to the fact these increments are $O(dt^{1/2})$ for a Wiener process. See section II of [19] for a pedagogical introduction.


[23] viz. the proof of Mallat's heuristic, i.e. that wavelet bases are optimal for representing functions containing singularities, in D. L. Donoho, Applied and Computational Harmonic Analysis 1, no.1, 100 (1993).


[25] cf. Eq. (3.11) with $m = 2$ in [6].

[26] viz. Eq. (B16) in [16].


[28] An interesting fact about the distribution of sizes of connected clusters is the existence of a critical black pixel probability $p_c \approx 0.5928$ such that an infinite (spanning) connected cluster occurs when $p > p_c$; see A. R. Conway, A. J. Guttmann, J. Phys. A 28, 891-904 (1995).


Bibliography


[16] See http://igec.inl.infn.it


[22] T. T. Lyons, M. W. Regehr and F. J. Raab, LIGO technical document P000014-00-D.


[34] A. Buonanno, Y. Chen and M. Vallisneri, gr-qc/0205122 (2002).


[47] Chernoff and Cordes, as cited in [43, section 4.1].


[62] The transfer function coefficients used were those calculated by Adhikari, Marka, Matone and Landry, http://blue.ligo-wa.caltech.edu/engrun/E7/Results/Calibration/index.html
