New Folder Name Detection, Measurement &
Gravitational Radiation

Detection, Measurement and Gravitational Radiation

Lee S. Finn
Department of Physics and Astronomy, Northwestern University, Evanston Illinois 60208
(20 July 1992)

Abstract

The optimum design, construction, and use of the Laser Interferometer Gravitational-wave Observatory (LIGO), the French/Italian gravitational wave observatory (VIRGO), or the Laser Gravitational-wave Observatory in Space (LAGOS) depends upon accurate calculations of their sensitivity to different sources of radiation. Here I examine how to determine the sensitivity of these instruments to sources of gravitational radiation by considering the process by which data are analyzed in a noisy detector. The problem of detection (is a signal present in the output of the detector?) is separated from that of measurement (what are the parameters that characterize the signal in the detector output?). By constructing the probability that the detector output is consistent with the presence of a signal, I show how to quantify the uncertainty that the output contains a signal and is not simply noise. Proceeding further, I construct the probability distribution that the parameterization $\boldsymbol{\mu}$ that characterizes the signal has a certain value. From the distribution and its mode I determine volumes $V(P)$ in parameter space such that $\boldsymbol{\mu} \in V(P)$ with probability P [owing to the random nature of the detector noise, the volumes $V(P)$ are are always different, even for identical signals in the detector output], thus quantifying the uncertainty in the estimation of the signal parameterization.

These techniques are suitable for analyzing the output of a noisy detector. If we are *designing* a detector, or determining the suitability of an existing detector for observing a new source, then we don't have detector output to analyze but are interested in the "most likely" response of the detector to a signal. I exploit the techniques just described to determine the "most likely" volumes $V(P)$ for detector output that would result in a parameter probability distribution with given mode.

Finally, as an example, I apply these techniques to determine the anticipated sensitivity of the LIGO and LAGOS detectors to the gravitational radiation from a perturbed Kerr black hole.

PACS numbers: 04.80.+z,04.30.+x,06.20.Dk,97.60.Lf

Typeset using REVTEX

I. INTRODUCTION

Under the present schedule, both the United States Laser Interferometer Gravitational Wave Observatory (LIGO [1,2]) and the French/Italian VIRGO [3] will begin operation in the late 1990s. Long before that time, theorists must lay a foundation for the study of gravitational radiation sources. Part of this foundation involves the construction of detailed, parameterized models of the waveforms from expected sources; another part involves the calculation of the anticipated sensitivity of the detector to each of these sources. Calculation of these kinds are not only needed for LIGO and VIRGO: design and technology studies for a Laser Gravitational-wave Observatory in Space (LAGOS) are currently being pursued [4] and calculations of the sensitivity of LAGOS to appropriate sources are needed to guide these studies.

In this paper I address the problem of calculating the anticipated sensitivity of a detector. like LIGO, VIRGO, or LAGOS, to an arbitrary source of gravitational radiation. The problem breaks up into two parts which I term *detection* and *measurement*. To "detect" is to decide whether the observed detector output contains a signal from a particular source or is just an example of noise; to "measure" is to assume the presence of a signal in the detector output and to characterize the signal in terms of the parameter(s) that describe the source (and its orientation with respect to the detector).

Echeverria [5] recently examined some of these issues in the particular context of determining the precision with which one could characterize the mass and angular momentum of a perturbed Kerr black hole from observations in a gravitational radiation detector. The foundation of his analysis was the construction of a quantity similar to the signal-to-noise ratio (SNR), and he asserted that the parameters that characterize a signal observed in the output of the detector are those that maximized this quantity. This analysis is limited in two respects (Echeverria & Finn [6]):

1. The validity of the formalism is restricted to the limit of high SNR; and

2. The formalism cannot determine the amplitude of the signal.

In addition, the conceptual basis of this calculation is not compelling: the determination of the parameters characterizing a signal in a noisy detector does not proceed by maximizing the SNR-like quantity defined by Echeverria [5].

In contrast, the techniques developed here are all based upon the construction of probabilities and probability densities. For the problem of detection, I construct the probability that the observed detector output is consistent with the presence (or absence) of a signal. In the case of the measurement problem, where the detector output is assumed to include a signal, the quantity constructed is the probability density that describes the likelihood of a given signal parameterization.

In §II I examine the twin processes of detection and measurement from the point of view of probability theory. The parameters characterizing a signal identified in the output of a noisy detector are defined to be those *most likely* to have resulted in the observed detector output. Some of the results described in this section are known elsewhere in the context of data analysis: they are included here for completeness and so that they may be compared with the techniques employed in Echeverria [5]. In §III I show how these same techniques can be exploited to evaluate the *anticipated sensitivity* of an instrument to a signal: *i.e.*, how precisely can the parameterization of a signal observed in the detector be determined. I find both exact and, in the interesting limit of a strong signal, approximate techniques for evaluating the expected precision with which an observed signal can be described. As an example, in §IV I apply the approximate techniques developed in §III to the determination of the parameterization of the gravitational radiation from a perturbed black hole, especially the black hole mass M and dimensionless angular momentum parameter a. In §V I briefly compare the methods and results of Echeverria [5] with my own. My conclusions are presented in §VI

II. DETECTION, MEASUREMENT AND PROBABILITY

In this section I consider two related problems that arise in the analysis of the output of a noisy detector: detection and measurement. The problem of detection is to determine whether or not a signal of known form (*i.e.*, deterministic, though parameterized by one or several unknown parameters) is present in the detector output. The problem of measurement is to determine the values of some or all of the unknown parameters that characterize the observed signal.

Note that the distinction between detection and measurement separates the determination of the presence or absence of the signal from the determination of the parameters that characterize it: detection does not address the value of the unknown parameters, and measurement presumes the signal's presence.

Detector noise can always conspire to appear as an example of the sought-for signal; alternatively, noise can mask the presence of a signal. In either case, noise interferes with our ability to determine the presence of the signal or the parameters that characterize it. Consequently, any claim of detection must be associated with a *probability* signifying the degree of certainty that the detected signal is not, in fact, an instance of noise. Similarly, when an observed signal is characterized it is appropriate to specify both a *range* of parameters and a *probability* that the signal parameters are in the given range.

For example, I can examine the data stream from a gravitational radiation detector to determine (with some probability) whether the radiation from the $l = |m| = 2$ mode of a perturbed, rotating black hole is present, irrespective of the black hole mass, angular momentum, or orientation with respect to the detector. If I conclude that a signal is present in the data stream, then I can attempt to determine bounds on some or all of these parameters, such that I expect the actual parameters characterizing the signal to fall within those bounds with a given probability.

In the next several subsections I examine detection and measurement in more detail. I assume that the statistical properties of the detector noise are known, and also that the *form* of the sought-for signal is known up to one or several parameters. My discussion focuses on determining the probability that a signal of known form is present in the output of a noisy detector, and on determining the probability that the unknown parameters have particular values.

While the discussion in §IV is framed in the context of the measurement of gravitational radiation from astrophysical sources, the questions addressed in this (and the following) se tion are purely statistical ones and contain nothing that is specific to gravitational radiation. general relativity, or any particular physical system or theory. For more details. the reader may consult Wainstein & Zubakov [7].

A. Detection

Consider a data stream $g(t)$ which represents the output of a detector. The data stream has a noise component $n(t)$ and in addition may have a signal component $m(t)$. The signal component is parameterized by several unknown parameters (denoted collectively as $\boldsymbol{\mu}$, and individually as μ_i); hence

$$g(t) = \begin{cases} n(t) & \text{if signal not present,} \\ n(t) + m(t;\boldsymbol{\mu}) & \text{if signal } m(t;\boldsymbol{\mu}) \text{ present.} \end{cases} \tag{2.1}$$

Assume that $\boldsymbol{\mu}$ is continuous, not discrete. I will describe how to determine the probability that $m(t;\boldsymbol{\mu})$, for *undetermined* $\boldsymbol{\mu}$, is present in $g(t)$, *i.e.*,

$$P(m|g) \equiv \begin{pmatrix} \text{The conditional probability that a} \\ \text{signal of the form } m(t;\boldsymbol{\mu}) \text{, for} \\ \text{unknown } \boldsymbol{\mu} \text{, is present given the} \\ \text{observed data stream } g(t). \end{pmatrix} . \tag{2.2}$$

Begin by using Baye's law of conditional probabilities to re-express $P(m|g)$ as

$$P(m|g) = \frac{P(g|m)P(m)}{P(g)}, \tag{2.3}$$

where

$$P(g|m) \equiv \begin{pmatrix} \text{The probability of measuring } g \\ \text{assuming the signal } m \text{ is present.} \end{pmatrix}, \tag{2.4a}$$

$$P(m) \equiv \begin{pmatrix} \text{The \textit{a priori} probability} \\ \text{that the signal } m \text{ is present.} \end{pmatrix}, \tag{2.4b}$$

$$P(g) \equiv \begin{pmatrix} \text{The probability that the} \\ \text{data stream } g(t) \text{ is observed.} \end{pmatrix}. \tag{2.4c}$$

Also re-express $P(g)$ in terms of the two possibilities m absent and m present, and further re-express the probability that m is present in terms of the probability that it is characterized by the *particular* $\boldsymbol{\mu}$:

$$\begin{aligned} P(g) &= P(g|0)P(0) + P(g|m)P(m) \\ &= P(g|0)P(0) + P(m)\int d^N\mu\, p(\boldsymbol{\mu})P[g|m(\boldsymbol{\mu})], \end{aligned} \tag{2.5}$$

where

$$P(0) \equiv \begin{pmatrix}\text{The } a\ priori \text{ probability that}\\ \text{the signal is } not \text{ present.}\end{pmatrix}, \tag{2.6a}$$

$$P(g|0) \equiv \begin{pmatrix}\text{The probability density of}\\ \text{observing } g(t) \text{ in the absence}\\ \text{of the signal.}\end{pmatrix}, \tag{2.6b}$$

$$P[g|m(\boldsymbol{\mu})] \equiv \begin{pmatrix}\text{The probability density of}\\ \text{observing } g(t) \text{ assuming}\\ m(t;\boldsymbol{\mu}) \text{ with } particular\\ \boldsymbol{\mu} \text{ is present.}\end{pmatrix}, \tag{2.6c}$$

$$p(\boldsymbol{\mu}) \equiv \begin{pmatrix}\text{The } a\ priori \text{ probability}\\ \text{density that } m(t) \text{ is}\\ \text{characterized by } \boldsymbol{\mu}.\end{pmatrix}. \tag{2.6d}$$

Combining equations 2.3 and 2.5, we find

$$P(m|g) = \frac{\Lambda}{\Lambda + P(0)/P(m)}, \tag{2.7}$$

where

$$\Lambda \equiv \int d^N\mu\, \Lambda(\boldsymbol{\mu}) \tag{2.8}$$

$$\Lambda(\boldsymbol{\mu}) \equiv p(\boldsymbol{\mu}) \frac{P[g|m(\boldsymbol{\mu})]}{P(g|0)}. \tag{2.9}$$

In equation 2.7 all of the dependence of $P(m|g)$ on the data stream g has been gathered into the *likelihood ratio* Λ. Aside from Λ, $P(m|g)$ depends only on the ratio of the *a priori* probabilities $P(0)$ and $P(m)$. In turn, the likelihood ratio depends on two components: the *a priori* probability density $p(m|\boldsymbol{\mu})$ and the ratio $P[g|m(\boldsymbol{\mu})]/P(g|0)$.

In order to determine $P(m|g)$ we must *assess* the *a priori* probabilities and *calculate* the likelihood ratio. It is often the case that we know, or can make an educated guess regarding, the *a priori* probabilities. For example, the sources may be Poisson distributed in time [determining $P(0)$ and $P(m)$], and they may be homogeneously distributed in space [determining $p(r) \propto r^2$, where r is the distance to the source]. At other times our assessment may be more subjective or based on imperfect knowledge, and in this case we can use the observed distribution of $\boldsymbol{\mu}$ to test the validity of our assessments using the techniques of hypothesis testing (Winkler [8] §7).

Now turn to the evaluation of $P[g|m(\boldsymbol{\mu})]/P(g|0)$. To determine this ratio, first note that the conditional probability of measuring $g(t)$ if the particular signal $m(t;\boldsymbol{\mu})$ is present is the same as the conditional probability of measuring $g'(t) = g(t) - m(t;\boldsymbol{\mu})$, assuming that the signal $m(t;\boldsymbol{\mu})$ is *not present* in g':

$$P[g|m(\boldsymbol{\mu})] = P[g - m(\boldsymbol{\mu})|0]. \tag{2.10}$$

Consequently, we can focus on the conditional probability of measuring a data stream $g(t)$ under the assumption that no signal is present [$P(g|0)$].

In the absence of the signal, $g(t)$ is simply an instance of $n(t)$. Assume that the $n(t)$ is a normal process with zero mean, characterized by the correlation function $C_n(\tau)$ [or, equivalently, by the one-sided power spectral density (PSD) $S_n(f)$]. In order to compute the ratio $P[g|m(\boldsymbol{\mu})]/P(g|0)$, consider the continuum limit of the case of discretely sampled data $\{g_i : i = 1, \ldots, N\}$, with the correspondence

$$g_i = g(t_i), \tag{2.11a}$$

$$t_i - t_j = (i-j)\Delta t, \tag{2.11b}$$

$$\Delta t = \frac{T}{N-1}. \tag{2.11c}$$

The probability that an individual g_i is a sampling of the random process $n(t)$ is given by

$$P(g_i|0) = \frac{\exp\left[-\frac{1}{2}\frac{g_i^2}{C_n(0)}\right]}{\left[2\pi C_n(0)\right]^{1/2}} \tag{2.12}$$

and the probability that the ordered set $\{g_i : i = 1, \ldots N\}$ is a sampling of $n(t)$ is

$$P(g|0) = \frac{\exp\left[-\frac{1}{2}\sum_{j,k=1}^{N} C_{jk}^{-1} g_j g_k\right]}{\left[(2\pi)^N \det ||C_{n,ij}||\right]^{1/2}} \tag{2.13}$$

where C_{jk}^{-1} is defined by

$$\delta_{jk} \equiv \sum_l C_{n,jl} C_{lk}^{-1} \tag{2.14}$$

and

$$C_{n,ij} \equiv C_n\left[(i-j)\,\Delta t\right] \tag{2.15}$$

(Mathews & Walker [9] §14-6, Wainstein & Zubakov [7] eqn. 31.11). Note that the normalization constant in the denominator of equation 2.13 is independent of the g_i; consequently, it does not affect the *ratio* $P[g|m(\boldsymbol{\mu})]/P(g|0)$. Since it is this ratio that we are interested in, without loss of generality drop the normalization constant from the following.

To evaluate equation 2.13 in the continuum limit, first note that

$$\delta(t_j - t_k) = \lim_{\substack{\Delta t \to 0 \\ T \to \infty}} \frac{1}{\Delta t}\delta_{jk}. \tag{2.16}$$

Consequently,

$$\begin{aligned} e^{2\pi i f t_k} &= \sum_j e^{2\pi i f t_j}\delta_{jk} \\ &= \lim_{\substack{\Delta t \to 0 \\ T \to \infty}} \frac{1}{\Delta t^2}\sum_j \Delta t\, e^{2\pi i f t_j} \sum_l \Delta t\, C_n(t_j - t_l) C^{-1}(t_l, t_k) \\ &= \lim_{\substack{\Delta t \to 0 \\ T \to \infty}} \frac{1}{\Delta t^2}\int_{-\infty}^{\infty} dt_j\, e^{2\pi i f t_j} \int_{-\infty}^{\infty} dt_l\, C_n(t_j - t_l) C^{-1}(t_l, t_k) \\ &= \lim_{\Delta t \to 0} \frac{1}{\Delta t^2}\int_{-\infty}^{\infty} dt_l\, e^{2\pi i f t_l} C^{-1}(t_l, t_k) \int_{-\infty}^{\infty} d\tau\, e^{2\pi i f \tau} C_n(\tau) && (2.17a) \\ &= \lim_{\Delta t \to 0} \frac{1}{\Delta t^2}\frac{1}{2} S_n(f)\widetilde{C^{-1}}(f, t_k). && (2.17b) \end{aligned}$$

To proceed from equation 2.17a to 2.17b, use the Wiener-Khintchine (*cf.* Kittel [10] §28) theorem to relate the PSD $S_n(f)$ to the correlation function $C_n(\tau)$ and define

$$\widetilde{C^{-1}}(f, t_k) \equiv \int_{-\infty}^{\infty} dt\, C^{-1}(t, t_k) e^{2\pi i f t}. \tag{2.18}$$

Consequently, as we approach the continuum limit, we have

$$\widetilde{C^{-1}}(f, t_k) = \lim_{\Delta t \to 0} \Delta t^2 \frac{2 e^{2\pi i f t_k}}{S_n(f)}. \tag{2.19}$$

With $\widetilde{C^{-1}}$ and Parseval's Theorem, we can evaluate the continuum limit of the argument of the exponential in equation 2.13:

$$\begin{aligned}
\lim_{\substack{\Delta t \to 0 \\ T \to \infty}} \sum_{j,k=1}^{N} C_{jk}^{-1} g_j g_k &= \lim_{\substack{\Delta t \to 0 \\ T \to \infty}} \frac{1}{\Delta t^2} \sum_{j,k=1}^{N} \Delta t^2 C^{-1}(t_j, t_k) g(t_j) g(t_k) \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t^2} \iint_{-\infty}^{\infty} dt_j\, dt_k\, C^{-1}(t_j, t_k) g(t_j) g(t_k) \\
&= \lim_{\Delta t \to 0} \frac{1}{\Delta t^2} \iint_{-\infty}^{\infty} df\, dt_k\, \widetilde{C^{-1}}(f, t_k) \tilde{g}^*(f) g(t_k) \\
&= 2 \int_{-\infty}^{\infty} df\, \frac{\tilde{g}^*(f)}{S_n(|f|)} \int_{-\infty}^{\infty} dt_k\, e^{2\pi i f t_k} g(t_k) \\
&= 2 \int_{-\infty}^{\infty} df\, \frac{\tilde{g}(f) \tilde{g}^*(f)}{S_n(|f|)}.
\end{aligned} \tag{2.20}$$

Here and henceforth we will denote the Fourier transform of $r(t)$ as $\tilde{r}(f)$.

Since the detector output $g(t)$ is real, $\tilde{g}^*(f) = \tilde{g}(-f)$. Define the symmetric inner product $\langle g, h \rangle$

$$\langle g, h \rangle \equiv \int_{-\infty}^{\infty} df\, \frac{\tilde{g}(f) \tilde{h}^*(f)}{S_n(|f|)}. \tag{2.21}$$

for real functions g and h. In terms of this inner product,

$$\begin{aligned}
\Lambda(\boldsymbol{\mu}) &= p(\boldsymbol{\mu}) \frac{P[g|m(\boldsymbol{\mu})]}{P(g|0)} \\
&= p(\boldsymbol{\mu}) \exp\left[2 \langle g, m(\boldsymbol{\mu}) \rangle - \langle m(\boldsymbol{\mu}), m(\boldsymbol{\mu}) \rangle\right],
\end{aligned} \tag{2.22}$$

The likelihood ratio Λ is found by substituting eqn. 2.22 into eqn. 2.8.

To summarize, the probability $P(m|g)$ that a signal of the class $m(t; \boldsymbol{\mu})$ is present in the output of the detector $g(t)$ can be expressed in terms of three *a priori* probabilities [$P(0)$, $P(m)$, and $p(\boldsymbol{\mu})$] and the ratio of two conditional probabilities [$P(m|g)/P(0|g)$]. The *a priori* probabilities must be assessed, while the ratio of the conditional probabilities can be calculated. Often we know, or can make an educated guess regarding, the *a priori* probabilities; at other times our assessment is subjective or otherwise based on imperfect knowledge. Finally we establish a threshold for $P(m|g)$ [or, equivalently, for Λ, $\ln\Lambda$, or some other surrogate of $P(m|g)$], and say that if the $P(m|g)$ (or its surrogate) exceeds the threshold then we have detected the signal.

I will not discuss detection further, except to say that the choice of threshold is influenced by our strategy to minimize errors. The two kinds of errors we can make are to claim the presence of a signal when one is in fact not present (a "false alarm"), or to dismiss an observed $g(t)$ as noise when a signal is present (a "false dismissal"). In order to *minimize* the probability of a false alarm (conventionally denoted α) we want a large threshold, while to minimize the probability of a false dismissal (conventionally denoted β) we want a small threshold. One obvious strategy for choosing the threshold is to minimize the sum $\alpha + \beta$, *i.e.*, to minimize the probability of making an error. Alternatively, some other combination of α and β may be minimized, taking into account the relative seriousness of the different kinds of errors. Regardless, it is inadvisable to blindly choose a threshold for $P(m|g)$ without careful consideration of the false alarm and false dismissal probabilities that arise and their relative severity.

B. Measurement

Turn now to the question of measurement. From equations 2.3, 2.5 and 2.9 we have

$$p[m(\boldsymbol{\mu})|g] = \begin{pmatrix}\text{The conditional probability that} \\ \text{the particular signal } m(t;\boldsymbol{\mu}) \text{ is} \\ \text{present in the data stream } g(t).\end{pmatrix} = \frac{\Lambda(\boldsymbol{\mu})}{\Lambda + P(0)/P(m)}. \qquad (2.23)$$

This conditional probability density is directly proportional to $\Lambda(\boldsymbol{\mu})$ and, since the denominator in equation 2.23 is independent of $\boldsymbol{\mu}$, it is maximized where $\Lambda(\boldsymbol{\mu})$ is maximized. If we assume that the signal is present, then the probability density that it is characterized by $\boldsymbol{\mu}$ is

$$p[m(\boldsymbol{\mu})|g,m] = \frac{\Lambda(\boldsymbol{\mu})}{\Lambda}. \qquad (2.24)$$

The goal of the measurement process is to determine a volume $V(P)$ in parameter space such that $\boldsymbol{\mu} \in V(P)$ with probability P. This volume is "centered" on the mode of the distribution $p[m(\boldsymbol{\mu})|g]$ in a way we define later on. The mode of either $p[m(\boldsymbol{\mu})|g]$ or $p[m(\boldsymbol{\mu})|g,m]$ is the $\boldsymbol{\mu}$ that maximizes $\Lambda(\boldsymbol{\mu})$. Denote the mode by $\widehat{\boldsymbol{\mu}}$.[1] While I will occasionally refer to $\widehat{\boldsymbol{\mu}}$ as the "measured" parameterization of the signal, bear in mind that $\widehat{\boldsymbol{\mu}}$ is only the *most likely* parameterization of the observed signal.

If we assume that the global maximum of $\Lambda(\boldsymbol{\mu})$ is also a local extremum, then $\widehat{\boldsymbol{\mu}}$ satisfies

---

[1]While we assume in what follows that the distribution has a single mode, the generalization to a multi-modal distribution is trivial.

$$0 = \frac{\partial \Lambda(\boldsymbol{\mu})}{\partial \mu_i}; \tag{2.25}$$

equivalently, $\widehat{\boldsymbol{\mu}}$ maximizes

$$\ln \Lambda(\boldsymbol{\mu}) = \ln p(\boldsymbol{\mu}) + 2\langle m(\boldsymbol{\mu}), g\rangle - \langle m(\boldsymbol{\mu}), m(\boldsymbol{\mu})\rangle , \tag{2.26}$$

i.e., it satisfies

$$0 = \frac{\partial \ln p(\widehat{\boldsymbol{\mu}})}{\partial \mu_i} + 2\left\langle \frac{\partial m}{\partial \mu_i}(\widehat{\boldsymbol{\mu}}), g - m(\widehat{\boldsymbol{\mu}})\right\rangle. \tag{2.27}$$

This final set of equations is in general non-linear and may be satisfied by several different $\boldsymbol{\mu}$. Some will represent local maxima, while others will correspond to local minima or inflection points; thus, equation 2.25 is a necessary but not sufficient condition for $\widehat{\boldsymbol{\mu}}$.

An important characterization of the strength of the signal in a detector is the signal-to-noise ratio (SNR). The "actual" SNR depends on the true parameterization of the signal $\widetilde{\boldsymbol{\mu}}$. We do not have access to $\widetilde{\boldsymbol{\mu}}$; however, we *do* know that the most likely value of $\widetilde{\boldsymbol{\mu}}$ is $\widehat{\boldsymbol{\mu}}$, and we define the SNR in terms of $\widehat{\boldsymbol{\mu}}$:

$$\rho^2 = 2\langle m(\widehat{\boldsymbol{\mu}}), m(\widehat{\boldsymbol{\mu}})\rangle . \tag{2.28}$$

The factor of two arises because the power spectral density $S_n(f)$ is one-sided while $\widetilde{m}(f)$ is two-sided. Note that ρ^2 is expressed in terms of the signal power (*i.e.*, it is proportional to the square of the signal amplitude). There is some ambiguity in the literature over whether "SNR" refers to ρ or ρ^2. We avoid the ambiguity by referring to either ρ or ρ^2 wherever the context demands it.

Having found the distribution $p[m(\boldsymbol{\mu})|g]$ (or $p[m(\boldsymbol{\mu})|g,m]$), we define the boundary of the volumes $V(P)$ to be its iso-surfaces. The probability P corresponding to the iso-surface $p[m(\boldsymbol{\mu})|g,m] = K^2$ is

$$P = \int_{p[m(\boldsymbol{\mu})|g,m]\geq K^2} d^N\mu\, p[m(\boldsymbol{\mu})|g]. \tag{2.29}$$

Note that since the distribution $p[m(\boldsymbol{\mu})|g,m]$ is not generally symmetric, $\widehat{\boldsymbol{\mu}}$ is not necessarily the *mean* of $\boldsymbol{\mu}$. Also, if the distribution $p[m(\boldsymbol{\mu})|g]$ has more than one local maximum then $V(P)$ need not be simply connected.

To summarize, suppose we have an observation $g(t)$ which we assume (or conclude) includes a signal $m(\widetilde{\boldsymbol{\mu}})$ (for unknown $\widetilde{\boldsymbol{\mu}}$). We construct the probability density $p[m(\boldsymbol{\mu})|g,m]$ according to equation 2.24, and identify iso-surfaces of $p[m(\boldsymbol{\mu})|g,m]$ as the boundary of probability volumes $V(P)$ according to equation 2.29. Finally, we assert that $\widetilde{\boldsymbol{\mu}} \in V(P)$ with probability P.

III. MEASUREMENT SENSITIVITY

In §II we saw how to decide whether a signal is present or absent from the output of a noisy detector, and, if present, how to determine bounds on the parameterization of the signal. Now I show how to *anticipate* the precision with which a detector can place bounds on the parameterization that characterizes a signal. In particular, consider an observed $g(t)$ which contains a signal $m(t;\tilde{\boldsymbol{\mu}})$ for unknown $\tilde{\boldsymbol{\mu}}$. We are interested ultimately in the distribution of

$$\delta\boldsymbol{\mu} \equiv \tilde{\boldsymbol{\mu}} - \hat{\boldsymbol{\mu}} \tag{3.1}$$

where $\hat{\boldsymbol{\mu}}$ is determined by the techniques discussed in §II. There are an infinity of possible $g(t)$ that can lead to the same $\hat{\boldsymbol{\mu}}$ [corresponding to different instances of the noise $n(t)$], and for each there is a different probability distribution $p[m(\boldsymbol{\mu})|g]$ (*cf.* eqn. 2.23) and a different set of probability volumes $V(P)$. We will find the probability volumes $V(P)$ corresponding to

$$p(\tilde{\boldsymbol{\mu}}|\hat{\boldsymbol{\mu}}) = \begin{pmatrix}\text{The conditional probability density}\\ \text{that the signal parameterization is}\\ \tilde{\boldsymbol{\mu}}\text{, assuming that the mode of the}\\ \text{distribution } p[m(\boldsymbol{\mu})|g] \text{ is } \hat{\boldsymbol{\mu}}.\end{pmatrix}. \tag{3.2}$$

I show first how to do this exactly, and then show a useful approximation for strong signals.

The mode $\hat{\boldsymbol{\mu}}$ of the distribution $p[m(\boldsymbol{\mu})|g,m]$ satisfies

$$2\left\langle m(\tilde{\boldsymbol{\mu}}) - m(\hat{\boldsymbol{\mu}}), \frac{\partial m}{\partial \mu_j}(\hat{\boldsymbol{\mu}})\right\rangle + \frac{\partial \ln p}{\partial \mu_j}(\hat{\boldsymbol{\mu}}) = -2\left\langle n, \frac{\partial m}{\partial \mu_j}(\hat{\boldsymbol{\mu}})\right\rangle \tag{3.3}$$

(*cf.* eqns. 2.1 and 2.27). Since $n(t)$ is a normal variable with zero mean, so are each of the $\langle n, \partial m/\partial\mu_j\rangle$ on the righthand side of equation 3.3. Denote these random variables ν_i:

$$\nu_i \equiv 2\left\langle n, \frac{\partial m}{\partial \mu_i}(\hat{\boldsymbol{\mu}})\right\rangle. \tag{3.4}$$

The joint distribution of the ν_i is a multivariate Gaussian and its properties determine, through the equation 3.3, the properties of the distribution of $\delta\boldsymbol{\mu}$. Consequently we can focus on the joint distribution of the ν_i.

Since the ν_i are normal, their distribution is determined completely by the means $\overline{\nu_i}$, which vanish, and the quadratic moments

$$\overline{\nu_i\nu_j} = 4\overline{\left\langle n, \frac{\partial m}{\partial \mu_i}(\hat{\boldsymbol{\mu}})\right\rangle \left\langle n, \frac{\partial m}{\partial \mu_j}(\hat{\boldsymbol{\mu}})\right\rangle}. \tag{3.5}$$

To evaluate the average on the righthand side, we will use the ergodic principle to turn the ensemble average over the random process n into a time average over a particular instance of n. Recalling that a time translation affects the Fourier transform of a function by a change in phase,

$$\mathcal{F}\left[r(t+\tau)\right] = e^{-2\pi i f\tau}\mathcal{F}\left[r(t)\right], \tag{3.6}$$

write

$$\langle n(t+\tau), r(t)\rangle = \int_{-\infty}^{\infty} dt\, e^{-2\pi i f\tau} \frac{\tilde{n}(f)\tilde{r}^*(f)}{S_n(f)}. \tag{3.7}$$

Consequently,

$$\begin{aligned}
\overline{\langle n,r\rangle\,\langle n,s\rangle} &= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} d\tau\ \langle n(t+\tau),r\rangle\,\langle n(t+\tau),s\rangle \\
&= \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} d\tau \int_{-\infty}^{\infty} df\, \frac{\tilde{n}(f)\tilde{r}^*(f)}{S_n(f)} e^{-2\pi i f\tau} \int_{-\infty}^{\infty} df'\, \frac{\tilde{n}(f')\tilde{s}^*(f')}{S_n(f')} e^{-2\pi i f'\tau} \\
&= \lim_{T\to\infty}\frac{1}{2T}\int_{-\infty}^{\infty} df\, \frac{\tilde{n}(f)\tilde{r}^*(f)}{S_n(f)} \int_{-\infty}^{\infty} df'\, \frac{\tilde{n}(f')\tilde{s}^*(f')}{S_n(f')} \delta(f+f') \\
&= \lim_{T\to\infty}\frac{1}{2T}\int_{-\infty}^{\infty} df\, \frac{\tilde{n}(f)\tilde{n}^*(f)}{S_n(f)} \frac{\tilde{r}^*(f)\tilde{s}^*(f)}{S_n(f)} && (3.8a) \\
&= \frac{1}{2}\int_{-\infty}^{\infty} df\, \frac{\tilde{r}^*(f)\tilde{s}^*(f)}{S_n(f)} && (3.8b) \\
&= \frac{1}{2}\langle r,s\rangle\,. && (3.8c)
\end{aligned}$$

In going from eqn. 3.8a to eqn. 3.8b, we used the definition of the PSD of the detector noise $n(t)$:

$$S_n(f) \equiv \lim_{T\to\infty}\frac{1}{T}\left|\tilde{n}(f)\right|^2 \tag{3.9}$$

(*cf.* Kittel [10] §28). With the result in equation 3.8c, we have

$$\begin{aligned}
\overline{\nu_i\nu_j} &= 4\overline{\left\langle n, \frac{\partial m}{\partial\mu_i}(\hat{\boldsymbol{\mu}})\right\rangle \left\langle n, \frac{\partial m}{\partial\mu_j}(\hat{\boldsymbol{\mu}})\right\rangle} \\
&= 2\left\langle \frac{\partial m}{\partial\mu_i}(\hat{\boldsymbol{\mu}}), \frac{\partial m}{\partial\mu_j}(\hat{\boldsymbol{\mu}})\right\rangle && (3.10a) \\
&\equiv \mathcal{C}_{ij}^{-1}. && (3.10b)
\end{aligned}$$

In terms of the $\mathcal{C}_{ij}$ (*i.e.*, the inverse of the $\mathcal{C}_{ij}^{-1}$), the joint distribution of the ν_i is given by

$$p(\boldsymbol{\nu}) = \frac{\exp\left[-\frac{1}{2}\sum_{i,j}\mathcal{C}_{ij}\nu_i\nu_j\right]}{\left[(2\pi)^N \det\left|\left|\mathcal{C}_{ij}^{-1}\right|\right|\right]^{1/2}} \tag{3.11}$$

This is also the joint distribution of the quantities

$$-2\left\langle m(\tilde{\boldsymbol{\mu}}) - m(\hat{\boldsymbol{\mu}}), \frac{\partial m}{\partial\mu_j}(\hat{\boldsymbol{\mu}})\right\rangle - \frac{\partial \ln p}{\partial\mu_j}(\hat{\boldsymbol{\mu}}) \tag{3.12}$$

that appear on the lefthand side of equation 3.3; consequently, we expect that for an observation characterized by a given $\hat{\boldsymbol{\mu}}$ the probability volumes $V(P)$ are given implicitly by

$$K^2 \geq \sum_{i,j} \mathcal{C}_{ij} \left[2 \left\langle m(\tilde{\boldsymbol{\mu}}) - m(\hat{\boldsymbol{\mu}}), \frac{\partial m}{\partial \mu_i}(\hat{\boldsymbol{\mu}}) \right\rangle + \frac{\partial \ln p}{\partial \mu_i}(\hat{\boldsymbol{\mu}}) \right] \times \left[2 \left\langle m(\tilde{\boldsymbol{\mu}}) - m(\hat{\boldsymbol{\mu}}), \frac{\partial m}{\partial \mu_j}(\hat{\boldsymbol{\mu}}) \right\rangle + \frac{\partial \ln p}{\partial \mu_j}(\hat{\boldsymbol{\mu}}) \right] \tag{3.13}$$

where

$$P = \int_{\sum_{i,j} \mathcal{C}_{ij}\nu_i\nu_j \leq K^2} d^N\nu \frac{\exp\left[-\frac{1}{2}\sum_{i,j} \mathcal{C}_{ij}\nu_i\nu_j\right]}{\left[(2\pi)^N \det ||\mathcal{C}_{ij}^{-1}||\right]^{1/2}}. \tag{3.14}$$

This result is exact as long as the maximum $\hat{\boldsymbol{\mu}}$ of $\Lambda(\boldsymbol{\mu})$ is also a local extremum of $\Lambda(\boldsymbol{\mu})$.

As the SNR becomes large the distribution $p(\tilde{\boldsymbol{\mu}}|\hat{\boldsymbol{\mu}})$ becomes sharply peaked about $\hat{\boldsymbol{\mu}}$ and the determination of the volume $V(P)$ is greatly simplified. Suppose that ρ^2 is so large that for $\tilde{\boldsymbol{\mu}} \in V(P)$ for all P of interest, the difference $m(\tilde{\boldsymbol{\mu}}) - m(\hat{\boldsymbol{\mu}})$ can be linearized in $\delta\boldsymbol{\mu}$. We then obtain in place of equation 3.3

$$\sum_i \delta\mu_i \mathcal{C}_{ij}^{-1} = -2\left\langle n, \frac{\partial m}{\partial \mu_j}(\hat{\boldsymbol{\mu}}) \right\rangle - \frac{\partial \ln p}{\partial \mu_j}(\hat{\boldsymbol{\mu}}) \tag{3.15}$$

The random variables $\delta\boldsymbol{\mu}$ are related to the $\boldsymbol{\nu}$ by a linear transformation,

$$\delta\mu_i = -\sum_j \mathcal{C}_{ij} \left[\nu_j + \frac{\partial \ln p}{\partial \mu_j}(\hat{\boldsymbol{\mu}})\right]; \tag{3.16}$$

consequently, the $\delta\boldsymbol{\mu}$ are normal with means

$$\overline{\delta\mu_i} = -\sum_j \mathcal{C}_{ij} \frac{\partial \ln p}{\partial \mu_j}(\hat{\boldsymbol{\mu}}), \tag{3.17}$$

and quadratic moments

$$\overline{\left(\delta\mu_i - \overline{\delta\mu_i}\right)\left(\delta\mu_j - \overline{\delta\mu_j}\right)} = \mathcal{C}_{ij}. \tag{3.18}$$

The probability distribution $p(\delta\boldsymbol{\mu}|\hat{\boldsymbol{\mu}})$ is a multivariate Gaussian (*cf.* eqn. 3.11):

$$p(\delta\boldsymbol{\mu}|\hat{\boldsymbol{\mu}}) = \frac{\exp\left[-\frac{1}{2}\sum_{i,j}\mathcal{C}_{ij}^{-1}\left(\delta\mu_i - \overline{\delta\mu_i}\right)\left(\delta\mu_j - \overline{\delta\mu_j}\right)\right]}{\left[(2\pi)^N \det ||\mathcal{C}_{ij}||\right]^{1/2}}. \tag{3.19}$$

Note that the matrix $\mathcal{C}_{ij}$ now has acquired a physical meaning: in particular, we see that the variances σ_i^2 of the $\delta\mu_i$ are

$$\begin{aligned} \sigma_i^2 &\equiv \overline{\left(\delta\mu_i - \overline{\delta\mu_i}\right)^2} \\ &= \mathcal{C}_{ii} \end{aligned} \tag{3.20}$$

and the correlation coefficients r_{ij} are given by

$$\begin{aligned} r_{ij} &\equiv \sigma_i^{-1}\sigma_j^{-1}\overline{\left(\delta\mu_i - \overline{\delta\mu_i}\right)\left(\delta\mu_j - \overline{\delta\mu_j}\right)} \\ &= \frac{\mathcal{C}_{ij}}{\sigma_i\sigma_j} \end{aligned} \tag{3.21}$$

In this sense we say that $\mathcal{C}_{ij}$ is the covariance matrix of the random variables $\delta\boldsymbol{\mu}$.

In the strong signal approximation, the surfaces bounding the volume $V(P)$ are ellipsoids defined by the equation

$$\sum_{i,j}\left(\delta\mu_i - \overline{\delta\mu_i}\right)\left(\delta\mu_j - \overline{\delta\mu_j}\right)\mathcal{C}_{ij}^{-1} = K^2, \tag{3.22}$$

where the constant K^2 is related to P by

$$P = \int_{\sum_{i,j}\mathcal{C}_{ij}^{-1}x^ix^j \le K^2} d^N x \, \frac{\exp\left[-\frac{1}{2}\sum_{i,j}\mathcal{C}_{ij}^{-1}x^ix^j\right]}{\left[(2\pi)^N \det\|\mathcal{C}_{ij}\|\right]^{1/2}} \tag{3.23}$$

It is often the case that not all of the parameters that characterize the signal are of physical interest. In that case, we may integrate the probability distribution (eqn. 3.11 or 3.19) over the uninteresting parameters, leaving a distribution describing just the parameters of physical interest.

Finally we come to the question of when the linearization in equation 3.15 is a reasonable approximation. Two considerations enter here:

1. It is important that the probability contours of interest (*e.g.*, 90%) do not involve $\delta\boldsymbol{\mu}$ so large that the linearization of $m(\widetilde{\boldsymbol{\mu}}) - m(\widehat{\boldsymbol{\mu}})$ is a poor approximation; and
2. It is important that the condition number (*cf.* Golub & Van Loan [11]) of the matrix $\mathcal{C}_{ij}^{-1}$ be sufficiently small that the inverse $\mathcal{C}_{ij}$ is insensitive to this approximation in the neighborhood of $\widehat{\boldsymbol{\mu}}$.[2]

These two conditions will depend on the problem addressed. If the validity of the linearization procedure is doubtful owing to the violation of either or both of these conditions, then we must fall-back on equation 3.3 and the exact results in equations 3.13 and 3.14.

IV. APPLICATION: A PERTURBED BLACK HOLE

In this section, I show how to use the approximate techniques developed in §III to find the precision with which the mass and angular momentum of a perturbed black hole can be

---

[2]Recall that the relative error in $\delta\boldsymbol{\mu}$ is the condition number times the relative error in $\mathcal{C}_{ij}^{-1}$: for a large condition number, small errors in $\mathcal{C}_{ij}^{-1}$ introduced by the linearization approximation can result in large errors in $\delta\boldsymbol{\mu}$.

determined through measurement in an interferometric gravitational wave detector. This problem was first considered by Echeverria [5].

Consider a single interferometric gravitational wave detector and a perturbed black hole of mass M and dimensionless angular momentum parameter a. Focus attention on a single oscillation mode of the black hole, *e.g.*, the $l = m = 2$ mode. The strain measured by the detector has the time dependence of an exponentially damped sinusoid characterized by the four parameters Q, f, V, and T:

$$h(t) = \begin{cases} 0 & \text{for } t < 0, \\ V^{-1/3} e^{-\pi f(t-T)/Q} \sin\left[2\pi f(t-T)\right] & \text{for } t > 0. \end{cases} \tag{4.1}$$

For convenience, assume that the perturbation begins abruptly at the *starting time* T. The *frequency* f depends inversely on the mass of the black hole, and has a weak dependence on its angular momentum: for the $l = m = 2$ quasi-normal oscillation mode,

$$f \simeq \frac{F(a)}{2\pi M} \tag{4.2a}$$

$$F(a) \equiv 1 - \frac{63}{100}(1-a)^{3/10}, \tag{4.2b}$$

is an accurate semi-empirical expression for the real part of the quasi-normal mode frequency (Echeverria [5] eqn. 4.4 and tbl. II). The *quality* Q is the damping time τ measured in units of the frequency f:

$$Q = \pi f \tau. \tag{4.3}$$

For the $l = m = 2$ oscillation mode of the black hole, Q depends entirely on a:

$$Q \simeq Q(a) \equiv 2(1-a)^{-9/20} \tag{4.4}$$

(Echeverria [5] eqn. 4.3 and tbl. II). Finally, the *amplitude* $V^{-1/3}$ of the waveform depends on the distance to the source, the size of the perturbation, and the relative orientation of the detector and the source.

This peculiar parameterization of the amplitude reflects our expectation that perturbed black holes are distributed uniformly throughout space (*i.e.*, $V \propto r^3$) and that all relative orientations of the detector and the black hole source are equally probable. Additionally, it reflects an assumption that perturbations of any allowed amplitude are equally probable; consequently, the *a priori* distribution $p(V)$ is uniform. Let us also assume that $p(a)$, $p(f)$ and $p(T)$ are uniform and that there is no *a priori* correlation of a, f, V, or T.

An interferometric gravitational wave detector is naturally a broadband receiver, though it can be operated in a narrow band mode (*cf.* Vinet, Meers, Man, & Brillet [12], Meers [13] and Krolak, Lobo & Meers [14]). Assume that the detector response function is uniform in the frequency domain over the bandwidth of the gravitational wave; consequently, the signal component in the output of the detector $[m(t, \boldsymbol{\mu})]$ is equal to the waveform $h(t; Q, f, V, T)$ (*cf.* eqn. 4.1). Assume also that the noise PSD (S_n) of the detector is independent of frequency (f) in the bandwidth ($1/\tau$) of the signal (I will discuss the validity of this approximation below).