COINCIDENCE ANALYSIS OF GRAVITATIONAL WAVE DATA

C.A. DICKSON

A thesis submitted to the University of Wales
for the degree of Doctor of Philosophy
September 1993

Declaration

I declare that this work has not already been accepted in any substance for any degree, and is not being currently submitted in candidature for any degree.

(Candidate)

Except where otherwise stated, this work is wholly the result of the candidate's own investigation. Suitable credit is given to joint work with colleagues, and to work of others throughout the thesis.

(Candidate)

(Supervisor)

Acknowledgements

I would like to thank my supervisor, Professor Bernard F. Schutz, for his guidance and countless helpful suggestions.

I would also like to acknowledge the help and support of all at the Department of Physics & Astronomy, University of Wales College of Cardiff; particularly Gabrielle Allen, Tim Barnett, Simon Chapman, Rob Davies, Dean Johnson, Gareth Jones, John Bryn Jones, David Nicholson, Justin Shuttleworth and John Watkins. Apologies to anyone whom I have missed.

Thanks are also due to many outside my institution who have assisted me, especially David Robertson at the University of Glasgow and Albrecht Rüdiger at the Max-Planck-Institut für Quantenoptik, Garching.

I acknowledge financial support from the S.E.R.C.; studentship number 8930551X.

To my Father and Mother

Abstract

The work presented herein falls into three parts. Part I reassesses the claims recently made by the Rome-Turin-Maryland (RTM) collaboration, that about the time of Supernova SN1987A, there were unusual correlations observed between four particle detectors and two room-temperature bar gravitational wave detectors. These correlations were claimed to have chance probability of as low as $\sim 10^{-6}$. By evaluation of RTM's *a posteriori* adjustment of many free parameters, I revise the probability estimates up to between $\sim 10^{-3}$ and the level of chance. I conclude, in contradiction to RTM, that the correlations are more likely due to chance fluctuations in the data than to a new physical effect.

Part II is a short, mainly discursive, section. Here, I state many lessons which can be learned from RTM's analysis, with particular relevance to the coincidence analysis which I perform in Part III.

In Part III, I perform the first coincidence analysis of data taken from interferometric gravitational wave detectors, the data coming from a coincident experiment lasting 100 hours (the *100 Hour Data Run*) in March 1989, between the prototype detectors at the University of Glasgow and the Max-Planck-Institut für Quantenoptik, Garching. In particular, I present the first working program for the coincidence analysis of data taken from two interferometric detectors. I devise efficient methods for vetoing untrustworthy data, including the *h-veto*, which removes coincidences which have measured amplitudes differing by more than a predetermined amount in probability space.

After applying these vetoes, I show that there were no highly improbable coincidences during the experiment. I place the first experimental limits on 10 kHz broadband gravitational waves: no coincidences were seen above $h = (6.8 \pm 1.3) \times 10^{-16}$ during the experiment. I also present a way in which this limit could be improved for future similar experiments. Finally, I list the lessons learned from my coincidence analysis, for the interest of experimenters and data analysts in the field.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Gravitational waves were theoretically predicted by Einstein in 1916, as a consequence of his general theory of relativity. Their existence is the most important remaining test of general relativity, and observation of gravitational waves will reveal much new astrophysics which will be interesting in its own right.

However, gravitational waves have not yet been convincingly detected, at the time of writing. Gravitational waves are very weak as a phenomenon; and at present, neither the detectors nor the data analysis systems are adequate to achieve reasonable observation rates. The solution of this detection problem and, later, the establishment of an observational science of gravitational wave astronomy, depends on attacking the problem from these two directions. On the one hand, one must build detectors which are sensitive enough. On the other hand, one must devise analysis methods and software to find the signals which may be there and, for the most part, to do this automatically and in real time.

With the next generation of detectors now being planned, we expect the increase in sensitivity obtained to facilitate the detection of gravitational waves by the end of the millenium. The overall system of data analysis also requires such an improvement. This thesis concerns itself with some of the remaining unsolved problems in the data analysis of gravitational waves. I hope it will contribute to the important first detection, when it happens; and, later, to the establishment of an observational science of gravitational wave astronomy.

1.1 Astrophysics of gravitational waves

1.1.1 Gravitational waves in general relativity

The field equations of general relativity, as derived by Einstein, are

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = kT_{\mu\nu}; \tag{1.1}$$

where $R_{\mu\nu}$ is the Ricci tensor, $R = R^{\mu}_{\mu} = g^{\mu\nu}R_{\mu\nu}$ is the trace of the Ricci tensor (the *Ricci scalar*), $g_{\mu\nu}$ is the metric tensor, $T_{\mu\nu}$ is the stress-energy tensor, and k is a constant. I use the following conventions:

- Strictly, $g_{\mu\nu}$ is not the metric tensor, but the (μ, ν) *component* of the metric tensor, normally denoted g; but I shall continue to use this lazy terminology. This also applies to the other "tensor" terms in the equation.
- The Greek indices μ, ν, etc. take values 0, 1, 2, 3. In flat space, I shall interpret these as the usual cartesian coordinates of special relativity, i.e. t, x, y, z, respectively.
- I use repetition of indices to indicate summation over the indices in question (the *summation convention*).

See, e.g. Schutz (1985) for more details. As is fairly standard for this calculation, I have ignored the cosmological constant term.

Consider the vacuum solution ($T_{\mu\nu} = 0$), and the weak field approximation, i.e.,

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}, \tag{1.2}$$

where $\eta_{\mu\nu}$ is the Galilean metric of flat spacetime, and $h_{\mu\nu}$ is a small perturbation, i.e. $|h_{\mu\nu}| << 1$. Now if we linearise and adopt the Lorentz gauge (see Schutz 1985), Eq. 1.1 becomes

$$\Box h_{\mu\nu} \equiv \left(-\frac{\partial^2}{\partial t^2} + \nabla^2\right) h_{\mu\nu} = 0. \tag{1.3}$$

This is d'Alembert's equation. The most simple solution is

$$h_{\mu\nu} = \mathrm{Re}\left[A_{\mu\nu}e^{2\pi i(t-z/c)}\right], \tag{1.4}$$

that is, the equation of a three dimensional wave propagating through spacetime in the z direction, at the speed of light, c. This is a gravitational wave. We shall return to their effect on matter, and how they are to be detected, in a moment. Firstly, a quick review of expected astrophysical sources of gravitational waves.

1.1.2 Astrophysical sources of gravitational waves

Gravitational waves are produced by all matter which is moving with a non-zero quadrupole moment. How much energy is released, and whether the waves are observable from sources at astrophysical distances, is another question. Even the most energetic sources predicted will be very difficult to observe with Earth-based detectors, because gravitational waves generally couple very weakly with matter. At present, there are four main expected sources of gravitational waves which we expect will be observable in ground-based detectors. These are the following (for more details, see, e.g. Thorne 1987).

Stellar collapse

At the end of its life, a star will suffer one or more collapses, because its radiation and gas pressures can no longer sustain the star against its own inward gravitational pull. As the core of the star collapses, it can give off gravitational radiation. Although only one supernova every hundred years or so is expected in a galaxy of our size, there may be many more collapses which are *electromagnetically-quiet*, i.e. do not have dramatic supernova-type optical or electromagnetic displays, or which are hidden in or behind dense gaseous clouds. Current guesses at event rates are of the order of one collapse per thirty years in our galaxy.

Of course, if one can observe out to more distant galaxies and other clusters of galaxies, the event rate will go up in proportion to the volume of space observed, i.e. to the cube of the distance out to which one can observe these phenomena. The prototype detectors working in Glasgow and Garching (see Section 1.2), and whose data I analyse in coincidence (see Part III), could only barely detect a nearby collapse event in our galaxy; while the next generation of long interferometric gravitational wave observatories (such as LIGO and VIRGO), with their much increased sensitivities, are expected to see collapse sources out to the Virgo Cluster, with an expected event rate of around several hundred per year (Hough *et al.* 1989).

Coalescing compact binary stars

All binary stars, due to the non-zero quadrupole moment of their orbit, are gradually losing energy in the form of gravitational waves. This will cause the orbit to decay[1], such that even-

[1] as is seen directly in the pulsar PSR 1913+16, whose orbital period is decreasing in exactly the way predicted by general relativity, to the observational limit of 1% of the effect (the effect being that the rate

tually the stars will coalesce. Calculations for compact binaries (see e.g. Schutz 1986; Thorne 1987 and references therein) show that most of the observable gravitational wave energy is given off in the last few seconds, when the orbital frequency reaches 100 Hz and above. (The close orbit of gaseous stars is much more difficult to simulate, due to hydrodynamics and tidal effects.)

Although the expected event rate for coalescing binaries in our galaxy is very low, compared to collapse event rates, the fact that one can fairly accurately predict the waveform of binary coalescence enables one to employ data analysis techniques, such as matched filtering, to improve the signal-to-noise ratio for a given detection; and hence to see objects much further away than would otherwise be the case. Hence, the observed event rate may be comparable to that of collapse events; but this is model-dependent (see Phinney 1991).

Continuous wave sources

Any rapidly rotating object will emit gravitational waves, if it has a non-zero quadrupole moment. Rapidly rotating neutron stars are the most famous candidate sources of this type of gravitational radiation. In order for them to emit, however, they must have some kind of non-axisymmetry (the axis in question being the rotation axis) due to either "geophysical" deformations (such as mountains) or large scale eccentricity of shape (e.g. caused by mechanical instabilities or magnetic effects). In this case, the expected wave form will be a sine wave, or several sine waves superimposed in the case of several non-axisymmetric imperfections in the shape of the body.

The expected amplitude of waves from such a source will be much lower again than for coalescing binaries; but the very long period of possible observation (up to years, interruptions being unimportant so long as the phase is preserved), coupled with our knowledge of the waveform, will enable observations of many sources in our galaxy with the long interferometers, operating in broadband. The main unsolved problem here in the data analysis is the inversion of the Doppler motion of the Earth, which is unknown *a priori* when the rotation frequency and location of the source on the sky are unknown. Some attempts have been made to solve this problem: see e.g. Schutz (1991).

of change of the orbital period, $\dot{P}_b = -2.43 \times 10^{-12}$ s s^{-1}). See Taylor & Weisberg 1989. This is almost conclusive observational evidence that gravitational waves exist, and are emitted at energies predicted by general relativity.

Stochastic background

Finally, it is expected that there will be other continuous sources which, although observable in amplitude terms, have unmodelled waveforms. These form the stochastic background of gravitational waves. Theoreticians have predicted that cosmic strings, if they exist, will be an observable source of gravitational waves. Furthermore, the Big Bang itself should be the source of an observable "echo" of gravitational waves (analogous to the 3 K cosmological microwave background). These waves are expected to be seen by the cross-correlation of the outputs of two or more detectors operating in coincidence over long periods.

1.2 Detectors of gravitational waves

Return now to Eq. 1.4. By imposing additional gauge conditions, and choosing the observers coordinate system, we can force the only non-zero components of $A_{\mu\nu}$ to be (see e.g. Schutz 1985),

$$A_{11} = -A_{22} \equiv A_+, \tag{1.5}$$

and

$$A_{12} = A_{21} \equiv A_\times. \tag{1.6}$$

Hence, from the geodesic differential equation (see Schutz 1985), one can show that if two particles are separated by ϵ in the x direction, the distance between them changes as the gravitational wave passes, such that

$$h_{xx} = -h_{yy} = \frac{2\delta\epsilon}{\epsilon}, \tag{1.7}$$

where $\delta\epsilon$ is the change in distance.

Thus, gravitational waves interact with matter by causing motion of particles in the plane perpendicular to that of wave propagation. The amplitudes of oscillation are dependent firstly on the amplitudes of the two *polarization* components, A_+ and $A_\times$, of the wave; and secondly on the distance between the bodies. The motion is such that, for example, if the wave has only the "+" polarization component, the space between two bodies in the x direction contracts as the space in the y direction expands, and vice versa.

This motion is the principle behind the two main types of gravitational wave detectors. These are as follows.

Laser interferometric gravitational wave detectors

The most common design for an interferometric gravitational wave detector is the Michelson interferometer. This exploits the motion induced in the two dimensions perpendicular to the direction of wave propagation. See Fig. 1.1.

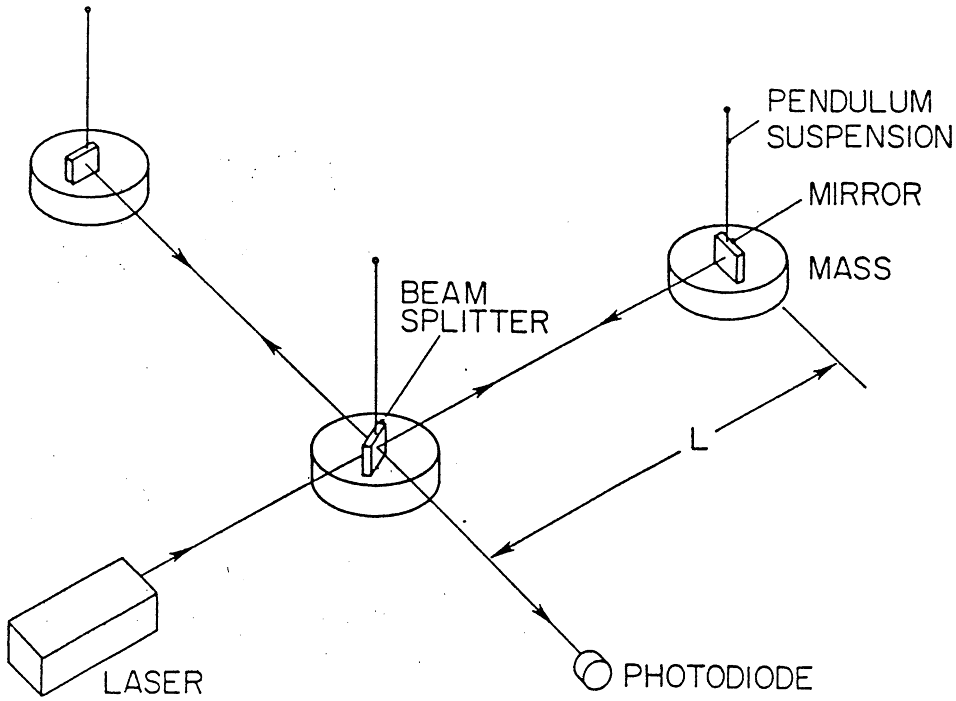


Figure 1.1: Basic design of a laser interferometric gravitational wave detector. Taken from Hough *et al.* (1986).

The mirrors at the ends of the arms reflect the laser light back to the beam splitter, where an interference pattern is observed. If a gravitational wave passes, the mirrors swing and the interference pattern changes. The effect is, however, very small. For a typical burst source in the Virgo cluster, with $h \sim 10^{-22}$, and with an effective arm length (many bounces of laser light are used in the arms of the interferometer, to increase the effective path distance between the test masses and the beam splitter) of 10^5 m, Eq. 1.7 gives the measured change in position of the masses as 10^{-17} m — smaller than the diameter of a proton.

Interferometric detectors are further subdivided into Fabry-Perot (resonant cavity) and delay line interferometers. I shall not go into more detail on these two types of interferometer: see, e.g., Hough *et al.* (1989).

At present, there are four working prototype laser interferometers in the world. These are situated at: the University of Glasgow; the Max-Planck-Institut für Quantenoptik, Garching, Germany; the California Institute of Technology; and at ISAS, Tokyo. The Glasgow prototype, for example, is of the Fabry-Perot design, while the Garching prototype is of the delay-line design.

Resonant bar detectors

Resonant bar detectors are slightly simpler in concept than are laser interferometers. A bar detector is a large cylinder of metal, usually aluminium alloy. If a gravitational wave passes the bar, at the appropriate frequency (the resonant frequency of the bar), it will excite the bar and cause it to "ring" at that frequency. A transducer converts this excitation to an output voltage. Again, the motions measured are very small; and thermal noise is the dominant noise source. Therefore modern bars are cooled down to cryogenic temperatures. About the time of the supernova SN1987A, typical bars functioned at room temperatures; while the new NAUTILUS ultracryogenic bar in Rome has recently been tested at below 100 mK (Astone *et al.* 1993).

The resonant frequencies of most bar detectors are around 1 kHz, about the expected frequency of the gravitational radiation emitted from collapse sources. This is also around the expected frequency of peak amplitude of gravitational wave emission from coalescing compact binaries.

Interferometers versus bars

For these two main types of detector, each has advantages over the other. The interferometer is broadband, i.e. is relatively sensitive to gravitational waves from frequencies of around 50 Hz right up to their sampling frequency, which may be 10 or 20 kHz. The bar is only sensitive in a narrow band (of the order of 2 Hz — Astone *et al.* 1993) about its resonant frequency. The interferometer doesn't need cooling as does the bar. Furthermore, the bar design seems to have a built-in sensitivity limit, given by the thermal quantum motions of the particles, at a sensitivity of around $h \sim 10^{-21}$ (Astone *et al.* 1993), which seems hard to beat no matter how low is the temperature. Unfortunately, this would seem to preclude the detection of bursts outside our own galaxy. The interferometer doesn't seem to suffer from this problem, in principle at least.

On the other hand, the bar is much cheaper and smaller than an interferometer of "comparable sensitivity" (if one can compare the sensitivities of a broadband and a narrowband detector). The interferometer requires expensive and bulky vacuum tubes, which the bar does not, and the interferometer arms would need to be hundreds of metres or even several kilometres long, in order to beat the bar in sensitivity at the bar's own resonance frequency (and to have reasonable event rates for detecting bursts and binary coalescences). The interferometer also requires expensive and advanced laser and mirror technology to operate at optimum sensitivity with arm lengths of the order of 3 km; which is the length scale on which LIGO are currently building and VIRGO are planning to build.

The basic choice between detectors of the interferometer or bar type is between broadband sensitivity and inexpensiveness; which is a matter of science versus funding. Of course, both types will be observing in the near future. The information they will provide will be complementary, and they may even be directly used together in coincidence (see e.g. Astone, Lobo & Schutz 1993).

1.3 State of affairs in gravitational wave research

Experimental gravitational wave astronomy as such began in the late 1950's with J. Weber, who constructed the first bar detector (Weber 1960). Of course, the sensitivity of the bar was far removed from what we expect are the sensitivity levels needed to observe gravitational waves from a distant source. Since then, many improvements have been made in bar technology (see e.g. Astone *et al.* 1993).

Prototype interferometers have also been built and run in the last twenty or so years. Again, great improvements have been made of late (see e.g. Newton 1993).

However, as I have said, there has not yet been a convincing claim for the detection of gravitational waves. There have been claims, but none of them has stood up to close scrutiny. We shall return to one such claim in Chapter 2, which at least involves bar detectors, even though the experimenters have not postulated a coherent model to explain what they see, nor even whether they believe that gravitational waves are actually involved.

Experimental gravitational waves, at least up to the time of writing, has been largely concerned firstly with setting better and better upper limits on sources; and secondly, in the case of interferometers, with improving the sensitivity of prototype detectors at various

Table 1.1: Collaborations to build long interferometric gravitational wave detectors

Collaboration	Institutions	Arm Length	Timescale[2]
LIGO	Caltech & MIT	2×4 km	1998
VIRGO	INFN Pisa & CNRS Orsay/Paris	3 km	?
GEO	U. Glasgow, MPQ Garching, & Cardiff U.	3 km ?	?
AIGO + collaborators	U.W. Australia, ANU, CSIRO Adelaide; IUCCA & CAT (India)	3 km	?
TENKO	ISAS (Tokyo)	100 m	1995

frequencies, and removing various noise sources in the detectors.

The task facing both the experimenters and data analysers is a huge one: much time and effort is being invested in solving all the problems associated with (1) the detection of the first gravitational wave, and (2) turning this nascent branch of astrophysics into a working observational science. There has been much collaborative effort amongst experimental teams, and between experimental teams and data analysis specialists: my own coincidence analysis of the joint 100 Hour Experiment is one of the fruits of the collaboration between the experimental groups at the University of Glasgow, Scotland and the Max-Planck-Institut für Quantenoptik in Garching, Munich, Germany, and the data analysis team at University of Wales College of Cardiff.

Long interferometers are very expensive and complicated to construct. Consequently, all but one of the current plans to build such detectors involve some sort of collaboration across institutions. At present, the main collaborative projects in the field are as shown in Table 1.1 (this is intended only as a rough guide, and in no way should be taken as a definitive statement on the subject).

The subject of collaborations has unavoidably become a sensitive political issue, so I will not dwell on this here. Note also that there are at least two proposed projects to put interferometers in space, namely LISA and SAGGITARIUS, but the funding and institutions involved in these are even more complicated.

The *status quo* in bar technology is simpler. The locations of the main bar detectors,

[2] Projected completion of building and preliminary operation.

along with a brief description of the types and temperatures of the bars, are given in Table 1.2.

Table 1.2: Existing gravitational wave bar detectors. (Based on Schutz 1989)

Institution	Room-temperature	Torsion	Cryogenic	Ultracryogenic
Stanford U.			√	√
Louisiana State U.			√	
U. of Maryland	√		√	
U. of Rome	√		√	√
U.W. Australia			√	
Moscow State U.	√			
Tokyo U.		√		
Guangzhou & Beijing	√			

The future for experimental gravitational wave research looks increasingly bright, with several long interferometers now at the planning stages, and with bars becoming more and more sensitive as they are cooled to lower temperatures. The data analysis problem, however, is a huge and underrated task in itself. It is with part of this problem that my thesis is concerned. The main topics I shall cover are as follows in the next section.

1.4 Topics covered in this thesis

Part I

The claims by the RTM (Rome-Turin-Maryland collaboration) groups, that there were strong correlations between various gravitational wave and particle detectors around the time of the supernova SN1987A in the Large Magellanic Cloud have stood unchallenged in the literature for some time. Indeed, RTM are still producing papers concerning this matter. However, many researchers believe that these correlations are more likely due to an artefact of RTM's data analysis than due to real signals of some sort, whether gravitational waves, neutrinos, some new particle or interaction, or a combination of these.

My doctoral supervisor and I have written a paper (Dickson & Schutz 1992) which performs a detailed reassessment on their analysis, complete with an attempt to account for every a *posteriori* choice or change of analysis parameters which we believe RTM used in order to optimise the correlations which they see. Our paper rebuts RTM's claims for very unlikely correlations, and we find that the correlations are very much more likely to have arisen by chance than RTM claim. Hence, we believe, there is no compelling evidence to conclude that gravitational waves or anything else have been seen in these data. A version of this paper is included verbatim as my Chapter 2. Although much collaboration with my supervisor was necessary in the writing of this paper, most of the original ideas are my own.

Following on from this is what could be described as an addendum to the paper contained in Chapter 2. This looks at some more claimed coincidences, similar to those in Chapter 2 and apparently occurring at the same time, between two neutrino detectors. This was not included in the original paper. It was written by me with very little discussion with my supervisor. The remainder of the thesis is exclusively my own work.

Part II

In Chapter 4, which forms the whole of Part II, I look at the lessons which can be learned from RTM's analysis of the gravitational wave and particle data. I find that they have a completely flawed approach to the analysis of these data, and draw various conclusions concerning the way I believe analysis of such data sets should and should not be performed. This generalises to other problems in physical sciences where the data have low signal-to-noise ratio, and where one may want to consider unmodelled or unexpected objects or even new physical processes as being candidate sources of signals in the data.

I use these lessons learned to make a short list of recommendations for future such analysis of gravitational wave data.

Part III

Bearing my own recommendations in mind, I then perform the first coincidence analysis of laser-interferometric data. Those data were taken from a coincidence experiment lasting 100 hours (the 100 Hour Experiment, or 100 Hour Data Run), in March 1989, between the prototype interferometers in Glasgow and Garching, to which I have already referred.

In Chapter 5 I review the background to the experiment, so the reader is familiar with the necessary details before I proceed.

Chapter 6 describes my methodology for the coincidence analysis, including my method for estimating the probability of any coincidences occurring by chance during the experiment.

I choose to concentrate on short duration collapse and burst sources for the coincidence analysis.

I tackle the large subject of *vetoes* in Chapter 7. This involves examining the various housekeeping data streams (seismometers, microphone signals, etc.) recorded in the laboratories of the two detectors, and using them to remove, or *veto*, events or coincidences which I deem untrustworthy. I also consider ways of removing coincidences which are untrustworthy for reasons other than that the housekeeping data were exhibiting unusual behaviour at certain times.

In Chapter 8 I give the results of my coincidence analysis. I set an upper limit on broadband bursts during the 100 Hour Experiment, and show that no highly improbable coincidences have occurred during the experiment.

My conclusions for the coincidence analysis are given in Chapter 9. I also list there the main achievements of my research in the coincidence analysis and in the writing of Part III.

At the end of my thesis I include two Appendices. Appendix A is a listing of my coincidence program, which was the first working program written to search for threshold-crossing coincidences in interferometric gravitational wave data. Appendix B is a list of my recommendations for future experiments, based on my experience with the coincidence analysis of the 100 Hour data. These recommendations have relevance to future experiments with other, larger interferometers, as well as to future experiments with these prototype detectors. They also pertain both to experiment and to data analysis, as these fields overlap considerably in practise.

Part I

Gravitational Wave – Neutrino Correlations

Chapter 2

Reassessment of the Reported Correlations between Gravitational Waves and Neutrinos Associated with SN1987A[1]

Abstract

Correlations of considerable apparent significance have been reported between data taken by two bar-type gravitational wave detectors and particle events recorded in the Mt. Blanc, Kamiokande, and IMB particle detectors during a 2-hour period near the explosion of supernova SN1987A. In particular, the correlations among the gravitational wave detectors and the Mt. Blanc neutrinos were claimed to have a chance probability of less than 10^{-6}. If this low probability implies that the correlations are a real physical effect, then new physics will be required to explain them. However, one of the statistical tests used to establish these correlations is seriously flawed, and most others were devised *a posteriori* and contain considerable freedom to make choices that affect the probability of finding correlations. By a careful consideration of these free parameters, and by applying similar analysis methods to

[1]Note: this chapter is a version of a paper written by C.A. Dickson and Bernard F. Schutz and submitted to Physical Review D in 1992

simulated pseudo-random data sets, we show that the actual frequency with which correlations similar to those in the Mt. Blanc data would occur in random data streams is between 0.1% and 10%. Moreover, if the Mt. Blanc correlations were real, then one would expect them in the other particle detectors. After inspecting the evidence, we also conclude that there are no physically significant correlations of the Mt. Blanc type between the gravitational wave detectors and the Kamiokande and/or IMB particles. This makes it very likely that the Mt. Blanc correlations are due, not to any physical effect, but simply to chance.

2.1 Introduction

At about the time of the supernova explosion SN1987A there were, unfortunately, only two gravitational wave detectors in operation (Amaldi *et al.* 1988). These were of the least sensitive type: room temperature bar detectors, one in Maryland and the other in Rome. There were four proton-decay experiments in operation that had the capability to detect particles from the supernova, and three of them — Kamiokande (Hirata *et al.* 1987), IMB (Bionta *et al.* 1987), and Baksan (Alexeyev *et al.* 1987) — registered a coincident burst. Unfortunately, only one gravitational wave detector was recording data at that time (Rome), and that was affected by seismic noise (Amaldi *et al.* 1988). However, at the time of a somewhat earlier "neutrino" burst in the Mont Blanc detector (Aglietta *et al.* 1987a), which probably was not associated with the supernova, both gravitational wave detectors were working satisfactorily.

Since gravitational waves emitted by the supernova and carrying any reasonable amount of energy would be well below the sensitivity limits of these room-temperature bar detectors, it was not expected that the gravitational wave data would show any signals. The first published analyses by the teams involved in the detection and analysis of the data, to whom we shall refer as the Rome-Turin-Maryland Collaboration (RTM) [2], found: (i) that with a delay of 1.4 s with respect to the 5 neutrino events of the apparent burst, the Rome gravitational wave data were at an appreciably higher level of excitation than average (in particular, there was an unusual excitation of the Rome detector just before the first Mont Blanc neutrino event (Amaldi *et al.* 1987), with a chance probability of 3%); and (ii) there

[2] Our use of the abbreviation "RTM" is a convenient shorthand for a large team of authors: not all individual authors contributed to all papers; and in two cases, Aglietta *et al.* (1989); Aglietta *et al.* (1991b), some authors from Moscow have contributed. See the individual citations for the full author lists.

was a modest correlation between the Rome and Maryland gravitational wave detectors in a 7 hour period around the time of the Mont Blanc burst (Amaldi *et al.* 1988) (with chance probability 3.5%) . But they reported no unusual coincidences between the two gravitational wave detectors just at the time of the Mont Blanc burst. On this evidence, there would be no reason to suppose that gravitational waves from SN1987A had been detected.

However, in subsequent analyses, RTM searched a larger stretch of data for further events like those reported earlier (Amaldi *et al.* 1987), where a gravitational wave detector is excited a fixed time before a particle is detected. This has led to a series of papers (Amaldi *et al.* 1987; Aglietta *et al.* 1989; Amaldi *et al.* 1989; Aglietta *et al.* 1991b; Aglietta *et al.* 1991a) reporting that time-delayed coincidences have occurred in various stretches of data with apparently high significance (low chance probability). RTM have found numbers of delayed coincidences between the gravitational wave detectors and the Mont Blanc neutrino detector (Amaldi *et al.* 1987; Aglietta *et al.* 1989) and between the gravitational wave detectors and the Kamiokande (Amaldi *et al.* 1989; Aglietta *et al.* 1991b), Baksan (Aglietta *et al.* 1991b), and IMB (Aglietta *et al.* 1991a) particle detectors respectively. RTM assigned chance probabilities to various of these coincidences in the range from 10^{-2} down to 10^{-6}. Our main purpose in this paper is to reassess these claims by RTM.

It seems clear that if the delayed coincidences are due to a real physical effect, then new physics will be required to explain them. Tens of coincident events are claimed to have taken place over a 2-hour period. If they are due to neutrinos and gravitational waves from SN1987A, the energy involved would be huge, many thousands of solar rest masses converted into gravitational wave energy *for each event* [3]. Moreover, given the low efficiency of neutrino detection, potentially thousands of events may have been missed. If they are not gravitational waves and neutrinos, then some new particles and interactions would be required.

One's attitude towards the need for new physics depends on (i) the *significance* of the observed correlations and (ii) one's assessment of the *plausibility* of the new physics required [4].

[3] The claim that gravitational wave detectors actually have a much larger cross-section than we have taken here — e.g., Weber (1984) — is wrong, as has been shown by Thorne (1992) and Grishchuk (1992); and so does not offer a way out of these problems.

[4] It would be interesting to take a more Bayesian approach to the whole analysis (see, e.g., Bulmer 1979), perhaps by assigning some *a priori* probability to the new physics required. Because RTM do not postulate any consistent physical explanation of their correlations, we shall not attempt this.

In practice, the significance of unusual results is usually taken to be the *a priori* probability of obtaining the results under a null hypothesis (normally that the data are completely random). The significance is particularly difficult to assess when, as here, the correlations were unexpected and so were only found *a posteriori*, after examination of at least part of the data set, and even then only by unusual statistical methods. RTM understand this and attempt to take account of the *a posteriori* nature of the effect by using other stretches of data as "control" sets in which to look for chance correlations. Unfortunately, we shall show that their "control" analysis is seriously compromised by the way the control data were chosen and by the statistical dependence of data they treat as independent. In fact, their principal statistical test is so flawed by data dependencies that we believe it is impossible to draw reliable conclusions from it.

We therefore undertake as part of our analysis to provide a more reliable control set by generating large numbers of random data sets on a computer and using RTM's own methods to analyze them. We directly address the question of how much freedom RTM had to find *a posteriori* effects in their original analysis of the Rome–Maryland–Mont Blanc data, such as by varying the time-delay and the thresholds of the gravitational wave detectors until they found significant coincidences. (RTM do in fact describe doing this.) Our approach cannot, of course, do more than estimate the true chance probability of the correlations, but it is a completely independent analysis, and it gives a radically different answer from the one RTM give.

Regardless of the significance of the Rome–Maryland–Mont Blanc correlations, the acid test of whether they point to a new physical effect is whether similar correlations occurred between the gravitational wave detectors and other particle detectors at the same time. RTM analyzed the data from the Kamiokande (Amaldi *et al.* 1989), Baksan (Aglietta *et al.* 1991b) and IMB (Aglietta *et al.* 1991a) detectors and claimed that they do in fact support the reality of the effect: they find correlations which they claim are very significant. Unfortunately, their analyses are again compromised by their data-selection criteria, by time-keeping problems in two of the detectors involved, and most seriously by the fact that, as we shall show, *RTM do not find significant correlations when they analyze the data in the same way as they analyzed the Mt. Blanc data.*

RTM themselves admit that, using these analysis techniques, there are no significant correlations between gravitational waves and Baksan particles (Aglietta *et al.* 1991b). They find modest correlations between gravitational wave data and IMB and Kamiokande particle events using the same analysis techniques, but we shall show that their analysis is fatally compromised by various data-selection criteria and by time-keeping problems. They find apparently significant correlations in these three detectors only when they use new methods of analysis not applied to the Mt. Blanc data. It is our conclusion that there is no evidence of correlations of the Mt. Blanc type in the Kamiokande, Baksan or IMB data, and that therefore the RTM correlations fail this crucial predictive test.

The data and analyses of RTM appear in a number of places in the scientific literature, some of them not widely available. We therefore shall try to make this paper as self-contained as possible. We begin in Section 2.2 with a review of the actual observations made by the two gravitational wave detectors and two particle detectors at the time of the supernova. In Section 2.3 we then give a summary of RTM's main analysis techniques. We point out that one of their unusual methods (which we call the net-excitation method) is seriously flawed. In Section 2.4 we present our own analysis of simulations of the Mt. Blanc and gravitational wave data streams, using mainly the other RTM method (the threshold-coincidence method). We find that the frequency distribution in random data sets of the sorts of correlations that RTM find is very much larger than RTM estimate. This allows us to make a detailed reassessment in Section 2.5 of the coincidence claims, including an attempt to correct for the large number of sometimes hidden degrees of freedom that have been used by RTM to optimize the correlations. These include the following:

1. *a posteriori* choices of, or freedom to choose, the time-delay

2. *a posteriori* choices of, or freedom to choose, the gravitational wave threshold

3. choice and variation of the duration of the data set

4. choice and variation of the starting time of the data set

5. statistical dependence of data sets caused by including the original "eyeballed" data set in the larger ones that were subjected to an analysis that was based on inspection of the original set

6. use of nonstandard and seriously flawed statistical tests with poorly-understood statistics, when standard tests could have been used but were not (or were not reported); and

7. the failure to apply consistently the Mt. Blanc analysis methods to data from Kamiokande and IMB.

(Some of the details of RTM's analysis are deferred to the appendix, with additional criticism where appropriate.)

The effects of some of these degrees of freedom are fairly easily quantified, while some are not so easily quantified. However, none of them is negligible; and all of them have the effect of making the coincidences more likely to have arisen by chance than RTM have claimed. Our reassessment for the gravitational wave—Mt. Blanc coincidences revises the coincidence probability from $\sim 10^{-6}$ (RTM's estimate) to 10^{-3}–10^{-1} (our estimate). For gravitational wave—Kamiokande coincidences we revise from $\sim 10^{-4}$ (RTM's estimate) to the level of chance (our estimate). Finally, for gravitational wave—IMB coincidences we revise from $\sim 10^{-3}$ (RTM's estimate) to $\sim 10^{-1}$ (our estimate). We feel that these correlations are therefore much more likely to have arisen by chance than to be a pointer to new physics.

RTM themselves never actually claim that the correlations are due to a real physical effect, and they have not proposed a serious model to explain them. They also remark in places that their probability estimates are only tentative in some respects. Their papers contain full descriptions of the tests that they report, which makes our reassessment possible. However, RTM themselves have not published a more detailed assessment of their significance estimates, and we wish to fill that gap here.

We wish to make clear at this point that it is not the goal of this paper to attempt to give a definitive set of rules of how we believe gravitational wave data *should* be analysed, which is a paper in itself, and which one of the authors will address in his thesis (CAD). However, we could make the following general recommendations:

- that all analyses of a data set, whether or not they give the results expected or desired by the analysers, should be stated;

- that data sets should be carefully examined individually and the results reported before they are combined;

- that the analysis methods used should be standard where possible, and that in any case the statistics of the analysis methods should be well understood or explained, and clear enough to be questioned easily;

- that a clear model should be given and tested (at least, the null hypothesis should always be tested);
- once a new model has been postulated on the basis of a given dataset, any new data should be analysed in the same way as the original data were.

2.2 The Gravitational Wave and Neutrino Observations

The observations of SN1987A are well documented (Trimble 1988), so we shall not review all of them here. However, we shall review the observations of the particle and gravitational wave detectors in operation at the time of the supernova.

Note that we have had some difficulty with our nomenclature, not knowing whether events crossing the threshold of a particle detector are neutrinos, some other particle, or random excitations in the detector (a normal background count); and this will vary from one detector to another. To call all the Mt. Blanc events *neutrinos*, for instance, would be presumptuous; and since RTM have still not provided a consistent model for the effect they see, we shall, where appropriate, enclose the word *neutrino* in quotes. For the other three particle detectors, we have generally used the word *particle*; though again, this should not be taken to imply that, in all cases, real particles have been detected, or that the particles are or are not neutrinos.

2.2.1 Particle observations

There were four particle detectors in operation during the relevant period: Mont Blanc (variously called UNO or LSD) (Aglietta *et al.* 1987a), Kamiokande (K II) (Hirata *et al.* 1987), IMB (Bionta *et al.* 1987) and Baksan (Alexeyev *et al.* 1987). All four were in operation during the whole of 22-23 February 1987. The optical brightening of the supernova took place between about 2h and 11h UT on 23 February. Neutrinos would have been expected at any time up to 24 hours before this, allowing time for the hydrodynamic shock to reach the star's surface and cause the optical display.

At about 2h 52m 37s UT, Mt. Blanc observed a burst of five "neutrino" events (Aglietta *et al.* 1987a). This burst had a probability about 2×10^{-3} of arising purely from the Poisson background during a period of 24 hours immediately preceding the observation of the optical supernova event (Aglietta *et al.* 1987a). However, this observation cannot easily be reconciled with those of the other detectors in operation since no significant particle bursts coincident with the Mt. Blanc event were observed in the other detectors. Therefore, the Mt. Blanc burst is usually distrusted (Trimble 1988).

The later burst, however, at about 7h 35m UT was certainly a real flux of neutrinos from the supernova: the other three particle detectors in operation all showed signals above the threshold levels about this time. Kamiokande (Hirata *et al.* 1987) detected 11 neutrinos at 7h 35m 35s UT (± 60 s) within a time interval of 13 s, with energies between 7.5 and 36 MeV. IMB (Bionta *et al.* 1987) reported 8 neutrinos at 7h 35m 41s UT during an interval of 6 s, with energies from 20 to 40 MeV. Baksan (Alexeyev *et al.* 1987) detected 5 neutrinos at 7h 36m 11s UT ($+2$ s, -54 s) during a time of 10 s, above an energy threshold of 12.0 MeV.

Mt. Blanc itself did not register an intrinsically significant burst at this time, although it did record two events at 7h 36m 00.5s UT and 7h 36m 18.9s, discovered in the off-line analysis (Aglietta *et al.* 1987b). This is not particularly worrying: since Mt. Blanc is smaller than KII and IMB, one would only have expected of the order of 1.5 neutrinos.

We have indicated above a very important point for our analysis, namely that two of the particle detectors had serious uncertainties in the offset of the experiment's clock relative to Universal Time: Kamiokande (Hirata *et al.* 1987) had an absolute timing uncertainty, Δt_K, of ± 60 s; while the absolute uncertainty Δt_B in the Baksan clock (Alexeyev *et al.* 1987) lay in the range $-54 \text{ s} < \Delta t_B < 2 \text{ s}$. The absolute timing of the other two detectors was more accurate, with Mt. Blanc (Aglietta *et al.* 1987a) accurate to ± 2 ms and IMB (Bionta *et al.* 1987) to ± 50 ms. The relative timing accuracy between particle events in any given detector was extremely good: the only uncertainty is the constant time shift between the detector clocks.

Given the fact that all three events were well above threshold and that the timing uncertainty allows them all to be coincident, there is little doubt that they are supernova neutrinos. However, the timing uncertainty makes it difficult to assess the probabilities of any coincidences between these detectors and the gravitational wave detectors. We shall return to this point in Sections 2.3 and 2.5.

2.2.2 Gravitational wave detectors

The Rome and Maryland room temperature bar gravitational wave detectors operated satisfactorily at least from 18h 24m 3s of 21 Feb 1987 to 6h 2m 3s of 23 Feb 1987, a period that includes the Mt. Blanc burst but excludes the time of the KII-IMB-Baksan events. Soon after 6h on 23 February, the Maryland detector experienced electrical problems; and at 7h 35m UT, the time of the KII-IMB-Baksan coincident burst, there were seismic disturbances in Rome. RTM confine all their analyses to the period before 6h 2m 3s on 23 February, when both gravitational wave detectors were working.

The Rome antenna has a mass of 2300 kg and a resonant frequency of 858 Hz. The Maryland antenna has a mass of 3100 kg and a resonant frequency of 1660 Hz.

The data sampling rate of the Rome detector was 1 Hz, while that of the Maryland detector was 10 Hz. In order to compare the two data sets, RTM averaged the Maryland data over 1 s intervals. This is 3 times longer than the optimum averaging time for this antenna, so the resulting data set has poorer than optimum signal-to-noise ratio by a factor of $\sqrt{3}$.

Before 6h, the gravitational wave detectors seem well-behaved. Events in both detectors followed fairly well an exponential (thermal) distribution in energy, although both detectors had some extra events at higher energies (Aglietta *et al.* 1989). RTM should, perhaps, have performed a more thorough investigation of the data from the individual detectors. The mean noise temperatures were approximately 28.6 K (Rome) and 29.8 K (Maryland).

The Maryland clock maintained an accuracy of ± 0.1 s during this period. The Rome clock did have an error, but careful study of its behavior after the end of the observation period led RTM to apply a correction of (-0.7 ± 0.1) s to obtain the true time.

2.3 Summary of the Main RTM Analysis Methods

Here we review the main methods of the RTM coincidence analysis.

2.3.1 The main RTM analysis methods

The RTM "net excitation" method

The first method is to sum the values of the combined gravitational wave streams at all "coincidence times", namely the arrival times of the "neutrinos" minus a fixed chosen time-delay. While this method is unusual, it is not necessarily implausible; however, its statistics are obscure. RTM assess the statistics by examining the behavior of their data set under simple modifications of the method, such as changing the time-delay. We shall see that there are serious difficulties with the manner in which they do this.

Calling the energy excitations of the Rome and Maryland antennae $E_R(t)$ and $E_M(t)$, respectively, the principal statistic used by RTM in their first analysis method is what we shall call the "net excitation" of the gravitational wave detectors over this period:

$$C_*(\phi) = \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} [E_R(t_i + \phi) * E_M(t_i + \phi)] , \tag{2.1}$$

where ϕ is a chosen offset time, t_i is the arrival time of the i-th "neutrino", N_ν is the total number of "neutrinos", and "$*$" indicates either "$+$" or "$\times$", depending on whether one is using the sum or product of the gravitational wave signals. When the offset ϕ is negative we shall refer to it as a *time-delay* (of the "neutrinos" relative to the gravitational waves), and an *advance* when positive. The values of $t_i + \phi$ are rounded to the nearest gravitational wave sampling time for the evaluation of E.

When adding the signals ("$*$" = "$+$"), one has to decide how to weight the two detectors. (This does not apply to the product algorithm, but most of RTM's analyses, including *all* their most improbable correlations (Aglietta *et al.* 1989; Amaldi *et al.* 1989; Aglietta *et al.* 1991b; Aglietta *et al.* 1991a) use the sum algorithm only.) The decision of RTM (Aglietta *et al.* 1989) is to normalize them by the mass of the detector, i.e., to divide the energy of the Maryland antenna by the ratio 3100/2300 of the mass of the Maryland detector to that of the Rome detector. This is somewhat arbitrary, since it takes no account of the large difference in the resonant frequencies of the two antennas, which implies that they respond to completely different parts of the spectrum of any gravitational wave event. Note that RTM also do *not* make any correction for the different orientations of the detectors.

RTM assess the significance of any result by comparing $C_*(\phi)$ with some "background" values of the same quantity, as determined by using different time delays in the two gravitational wave streams:

$$C_*(\delta_1, \delta_2) = \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} [E_R(t_i + \delta_1) * E_M(t_i + \delta_2)] , \tag{2.2}$$

where δ_1 and δ_2 are separate time delays. We shall see in a moment that this definition of a comparison background fatally flaws this method.

By changing the time delays, RTM calculate a large number N_b of these background values, between $N_b = 10^3$ and $N_b = 10^6$ in various investigations. They then assign a ranking order to the various time-delays ϕ by defining

$$n(\phi) = \mathrm{count}_{\delta_1,\delta_2}\left[C_*(\delta_1,\,\delta_2) \geq C_*(\phi)\right], \tag{2.3}$$

where "$\mathrm{count}_{\mathbf{range}}[<\text{condition}>]$" means that one counts the number of times the condition is true for variables in the given range. In this case, the smaller is the value of $n(\phi)$, the more significant is the correlation for that time-delay. Since the range of δ_1 and δ_2 always includes ϕ, the minimum value of $n(\phi)$ is 1. The maximum is N_b.

On the assumption that the background values are all independent, the probability of the correlation at a given delay is then taken by RTM to be

$$p(\phi) = n(\phi)/N_b. \tag{2.4}$$

If the background values $C_*(\delta_1,\,\delta_2)$ were all independent and had the same distribution as the values of $C_*(\phi)$, and if $n(\phi) \gg 1$, then this would not be an unreasonable way of estimating the probability. Unfortunately, none of these three conditions holds in the RTM analysis. We shall examine the independence of the background values in a moment. (We discuss the effect of small-number statistics $[n(\phi) \sim 1]$ in Appendix 2.7.1, and we return to the question of the distributions of $C_*(\phi)$ and $C_*(\delta_1,\,\delta_2)$ in Appendix 2.7.2.)

Notice that this method uses only the ranking order of the values of the correlations, and does not attempt to use a frequency distribution in uniform steps of C_*, which would be more conventional. This means that two values of ϕ may give values of C_* that are very close, but they could be far apart in $n(\phi)$.

Criticism of the net excitation method

The biggest problem with the net excitation method is that the background values are not all independent. This is easy to see if we count the number of data points from which the background values are derived. RTM say that they always take values of δ_i such that the background value is taken from the same period as the signal, $C(\phi)$ (Aglietta *et al.* 1989). This is to avoid problems due to possible nonstationarity of the noise. Now, in a 2-hour stretch of data, where RTM find their strongest correlations (Aglietta *et al.* 1989), there are 7200 1-second samples from each gravitational wave detector. On the null hypothesis (no genuine correlation), there are thus about 1.4×10^4 independent random numbers in the original data. These numbers are combined in various ways using Eq. 2.2 to form up to 10^6

background values. There must, therefore, be hidden correlations among the background values, at least when N_b exceeds about 10^4. It would not be easy to characterize these correlations, but it would be most unwise to assume, as RTM do, that there are none of significance for this method. Any estimate of probability from this method below a few times 10^{-4} cannot, therefore, be reliable.

Indeed, we shall see that the results of this test, as reported by RTM, show great variations in the apparent probabilities for time-delays separated by as little as 0.1 s, well below the physical resolution of the gravitational wave experiments. This may well be due to the untrustworthiness of Eq. 2.4 for the smallest apparent probabilities.

Threshold coincidence method

The second RTM method is similar to the threshold-crossing gravitational wave–neutrino method we suggested at the beginning of this section, only it is applied to the combined gravitational wave data stream rather than to each one separately. RTM set a threshold on the combined data stream

$$E_*(t) = E_R(t) * E_M(t) \tag{2.5}$$

(where again "$*$" is "$+$" or "$\times$"), and identify gravitational wave "events" as those which cross the threshold. (These are not of course necessarily real gravitational waves: they may be just thermal noise excitations.) A coincidence occurs for a time-offset ϕ with a "neutrino" that arrived at time t if the event occurs at the nearest gravitational wave sampling time to $t + \phi$. The statistics of this method are much more straightforward, at least for a fixed threshold.

For a data set lasting N_t sampling intervals (of one second), containing N_ν "neutrinos" and N_{gw} gravitational wave events randomly (uniformly) distributed, the expected number of coincidences is

$$\bar{n} = \frac{N_\nu N_{gw}}{N_t}. \tag{2.6}$$

Given that arrival times are uniformly distributed, the probability of obtaining n or more coincidences, given the mean $\bar{n}$, is

$$p_{\bar{n}}(n) = \sum_{r=n}^{\infty} \frac{\bar{n}^r e^{-\bar{n}}}{r!} = 1 - \sum_{r=0}^{n-1} \frac{\bar{n}^r e^{-\bar{n}}}{r!}. \tag{2.7}$$

This equation holds provided $|\phi|$ is much less than the expected interval between coincidences; if $|\phi|$ is too large, end effects will reduce the coincidence probability.

2.4 Monte Carlo simulations

2.4.1 Computer model

The objective of our Monte Carlo computer simulation was to assess the realistic probability that the correlations found by RTM would arise by chance in completely random data sets. With computer-generated data we can experiment with changing thresholds, time delays, and even methods of analysis to see what effect these have on apparent correlations. We have generated large numbers of pseudo-random data, analyzed them using the RTM threshold-coincidence method, computed the apparent probability of the strongest correlations by RTM's net-excitation method, and then compared this with the actual relative frequency of occurrence of such correlations among the pseudo-random data sets. We principally simulate the analysis of the Mt. Blanc data, although our results also illuminate the treatment of the Kamiokande and IMB data.

2.4.2 Properties of the pseudo-random data

In each Monte Carlo run, two sets of artificial gravitational wave data were generated, one corresponding to the energy excitation of the Maryland detector and the other to that of the Rome detector. Each artificial data set consisted of 7200 samples, equivalent to a 2-hour data stream sampled at 1 s intervals. The samples were drawn from distributions which were exponential in the temperature of the excitation, the Rome simulated data with mean 28.6 K and the Maryland with mean 22.1 K (its effective temperature after normalizing its mass to that of the Rome antenna and averaging over 1 s intervals for comparison with the 1 Hz Rome data (Aglietta *et al.* 1989)).

For the "neutrinos", we assumed an exponential distribution of the time delays between one neutrino and the next, using the observed mean arrival interval in the Mt. Blanc data (Aglietta *et al.* 1989). (This is the distribution one expects, of course, if the neutrinos arrive according to the standard Poissonian "shot noise" model.)

To generate the random numbers we used the *Numerical Recipes* (Press *et al.* 1988) uniform random number algorithm *RAN1*. The cycle length of this random number generator is said to be infinite for all practical purposes (Press *et al.* 1988). We demonstrate its distribution by generating and binning the first 4×10^5 numbers in Fig. 2.1.

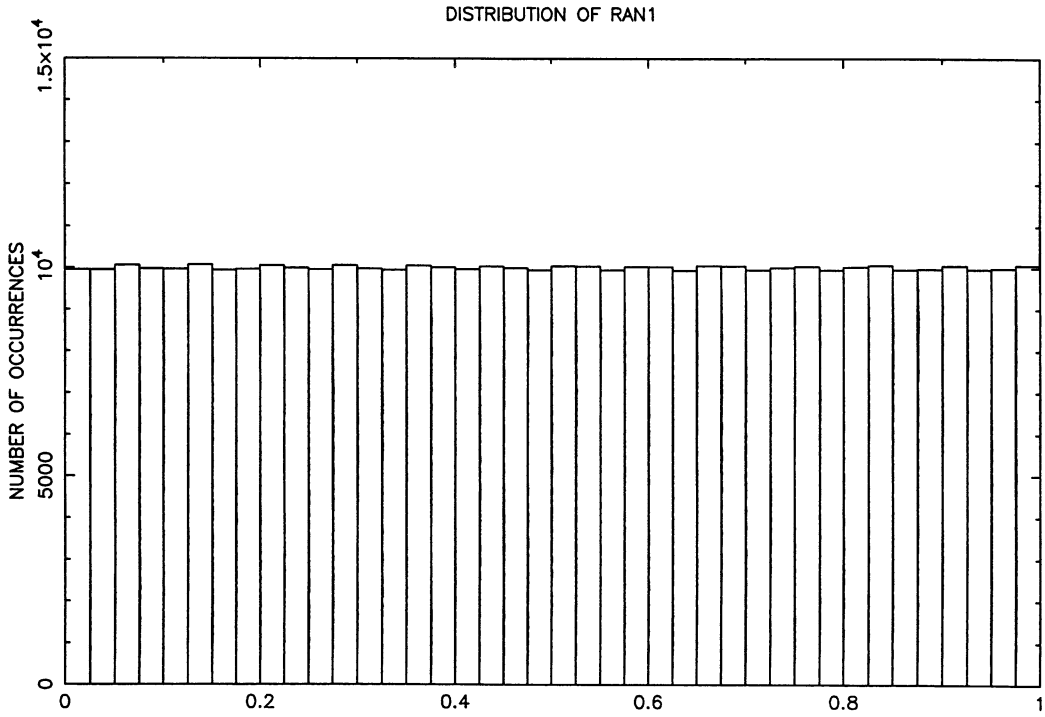


Figure 2.1: Distribution of the pseudo-random number generator

2.4.3 Method of analysis of the pseudo-random data

Adoption of the threshold-coincidence method

To analyze the simulated data one needs to choose from the large variety of statistical tests which RTM employ. Since the main thrust of this paper is to examine the Mt. Blanc-gravitational wave coincidences, we shall perform an RTM-style analysis of two artificial gravitational wave data streams and one artificial neutrino stream.

We perform an RTM-style threshold-coincidence analysis on each random set. This allows us to assess the influence of the freedom to choose the best threshold on the chance of finding a strong correlation. In view of the dubious value of the net-excitation measure of correlations, it would be inappropriate to subject each random set to such an analysis. Indeed, the computer time that would be required for such an analysis would be huge, since millions of random numbers would be required for the analysis of each data set. (Once a given set is generated, one needs to generate from it all the background values.) Instead, only for any data sets in which we found significant threshold-correlations do we also perform a net-excitation analysis. We will see that this still sheds considerable light on the question of how unusual are the correlations whose claimed (apparent) probability is 10^{-6}.

For each Monte Carlo data set, we have two choices to make, the threshold and the time-delay. We shall discuss each of these choices in turn.

Selection of a threshold

In choosing the threshold T, we are guided by what RTM say about their choice (Aglietta *et al.* 1989). They select $T = 150$ K for the summation statistic in the net-excitation method because it gives the best correlation. They indicate that they searched thresholds from 100 K to 200 K in steps of 10 K. In our simulations, therefore, we search through the same set of thresholds. This is a minimal set: we can be confident they searched all of these. If in fact they searched a larger number than they displayed in Fig. 16 of Aglietta *et al.* (1989), then the "true" probability of a correlation would be larger.

Choosing a time delay

Although it is clear that RTM searched some range of time delays before settling on their preferred one of 1.1 s, the central problem for us is to decide how wide that range should be when we analyze our simulated data. Note that, despite our reservations about the wisdom of varying time-delays in steps of only 0.1 s when the gravitational wave data have a time-resolution of 1 s, we must follow RTM in this if we are to simulate their methods faithfully.

In analysing the Mt. Blanc data, RTM changed their "best" delay from 1.4 s (Amaldi *et al.* 1987) to 1.2 s (Aglietta *et al.* 1989) and then to 1.1 s (Aglietta *et al.* 1989), depending which was the optimum delay for the data under consideration and the analysis method in question, so some *a posteriori* adjustments were made. RTM thus indicated their willingness to optimise the time delay, within a not-well-defined range, on receipt of more data and the use of other analysis methods. (In the case of the net excitation in Aglietta *et al.* (1989), this optimisation changes the "probability" from 10^{-4}, at delay 1.4 s, to 10^{-6}, at 1.1 s: a large change in "probability" for an apparently insignificant change in delay.) However, RTM never went far from their first value of 1.4 s, which they found by inspection from the raw Rome and Mt. Blanc data at about the time of the 5ν burst. RTM also tell us that they would never have adjusted the delay by more than about 0.5 s–1 s (Pizzella & Pallottino 1991), although of course this comment was made after publishing the results.

The crucial question for us is the following. Given that the initial eyeballing of the data had provided the motivation to search time-delays, if it then happened that, after receipt of the Maryland data and a full analysis of both gravitational wave datasets, RTM had discovered a much stronger correlation at a very different delay, would they have ignored it? Would they have been bound by their original choice of $1.4\pm$ (say) 0.5s when the phenomenon, by hypothesis, occurs over a period of 2 hours in both data sets, and when their original choice was made simply by crude eyeball inspection of 50 s of one of the data sets? We believe that, had a much better delay been found, RTM *should* have rejected their original choice completely.

Moreover, RTM *did* in fact search a wide range of time-delays after receiving the Maryland data. Using the net excitation method, they looked at delays from -3.2 s to $+0.8$ s, which was not necessary for the calculation of the strength of the correlation at -1.4 s, which they had postulated. It was this search that led to their later adoption of a delay of 1.1 s. Also, using the threshold coincidence method, RTM searched from -50 s to $+50$ s, for a fixed threshold, and found no correlations stronger than those around 1.2 s. This is not surprising since the threshold was optimised for the chosen delay of 1.2 s. If, in either of these searches, they had found any correlations which were stronger yet, RTM would surely have been obliged to take them seriously.

We conclude, therefore, that we should search our simulated data sets over a wide range of delays. This view is reinforced by an examination of RTM's initial selection of a 1.4 s delay.

On the initial selection of the 1.4 s time-delay. RTM initially inspected a small stretch of data containing the 5 Mt. Blanc "neutrinos" (see our Fig. 2.2; only the "neutrinos" and the Rome data were used), and they chose a delay for which the gravitational waves are "in most cases appreciably higher than the average background" (Amaldi *et al.* 1987). Since this criterion is just an "eyeball" implementation of their own net excitation method (Eq. 2.2) adapted for one detector instead of two, we shall now use this method to attempt to quantify the effect of their inspection process.

The first RTM time-delay estimate involved only the Rome data, so in Fig. 2.3 we plot the statistic

$$C(\phi) = \frac{1}{N_\nu} \sum_{i=1}^{N_\nu} E_R(t_i + \phi), \qquad (2.8)$$

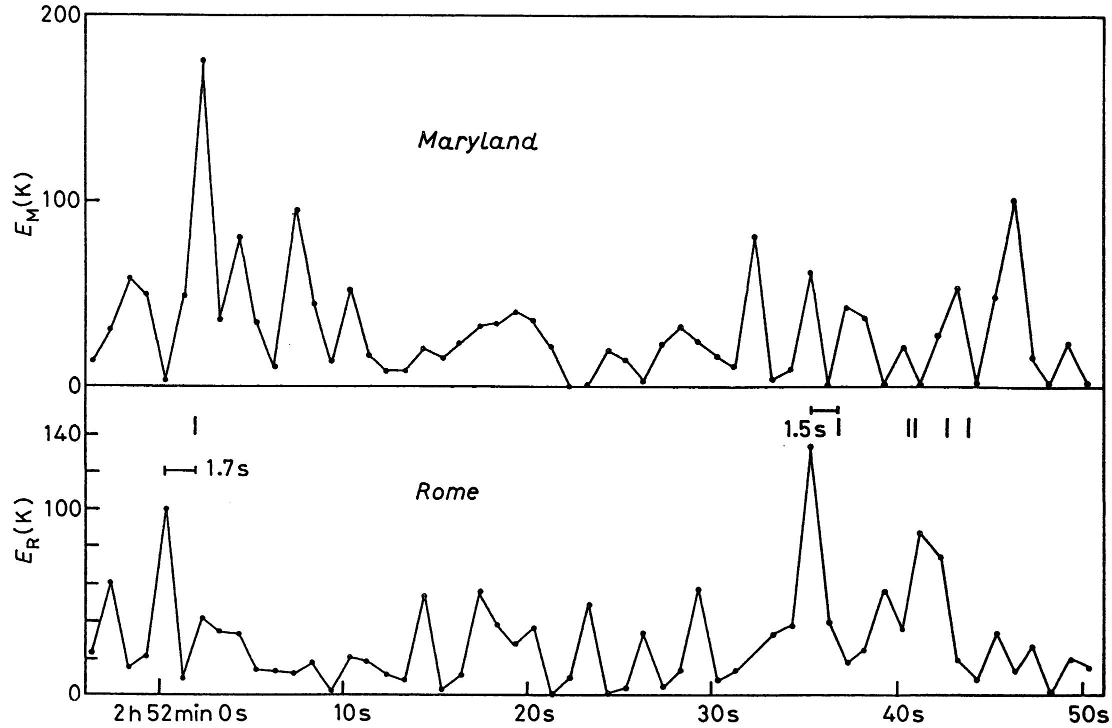
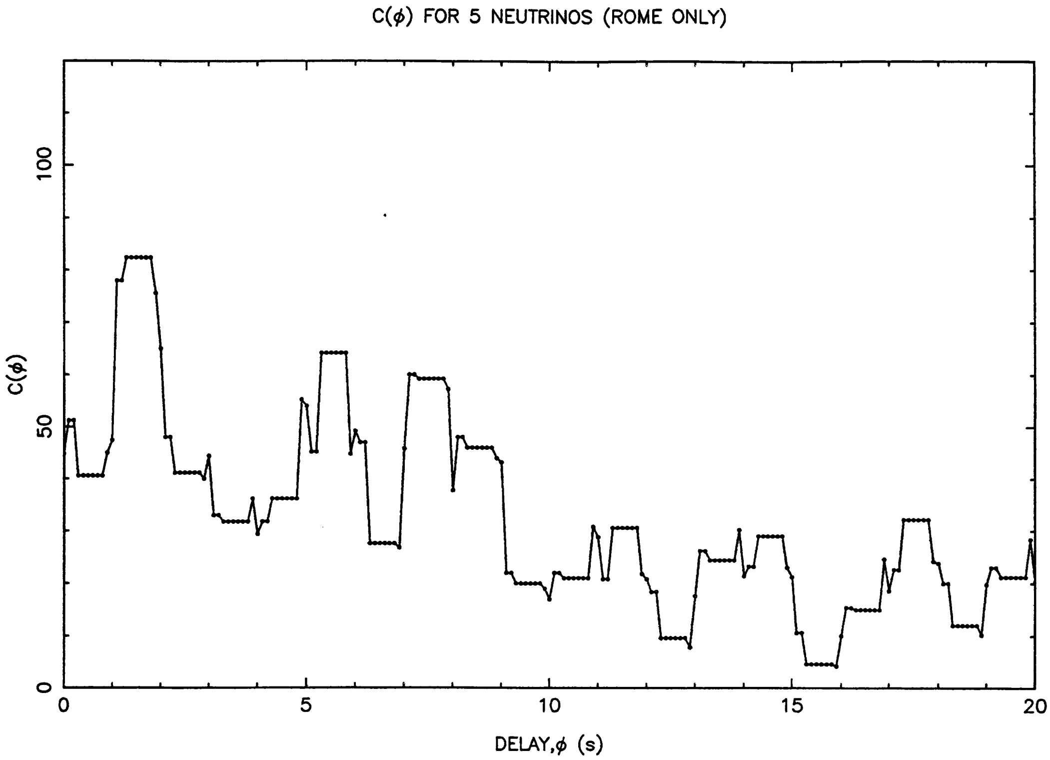


Figure 2.2: First indications of the Mt. Blanc correlations. RTM originally had only the Mt. Blanc "neutrinos" and the Rome data from which to select a delay of 1.4 s. The Maryland data were obtained later, and appear in the analyses in Amaldi *et al.* 1988. (Reproduced from Aglietta *et al.* 1989 with permission.)

which is the single-detector version of C_* of Eq. 2.1. Our figure contains two plots: (a) uses all six "neutrinos" that are shown in Fig. 2.2; (b) uses only the five "neutrinos" of the Mt. Blanc burst. In both cases, the best time-delays are between 1.3 s and 1.8 s, but there is no preference among them. This agrees with the RTM choice. But other delays offer hope of some effect: near 5.5 s and 7.5 s there are peaks above 50 K. Note that each peak is about 1 s wide, which agrees with the time-resolution of the gravitational wave data. When RTM broadened the analysis from the Rome–Mt. Blanc to the Rome–Maryland–Mt. Blanc data, they changed the delay from 1.4 s to 1.2 s after a similar eyeball inspection of a short stretch of the data (Amaldi *et al.* 1987; Aglietta *et al.* 1989). Accordingly, we next look at the effect on the delay when we include the Maryland data. Thus, we next look at the full net-excitation statistic C_+ (Eq. 2.1) applied to all the data of Fig. 2.2. Our results are shown in Fig. 2.4. Here, the picture is very different: there is little to choose between time delays



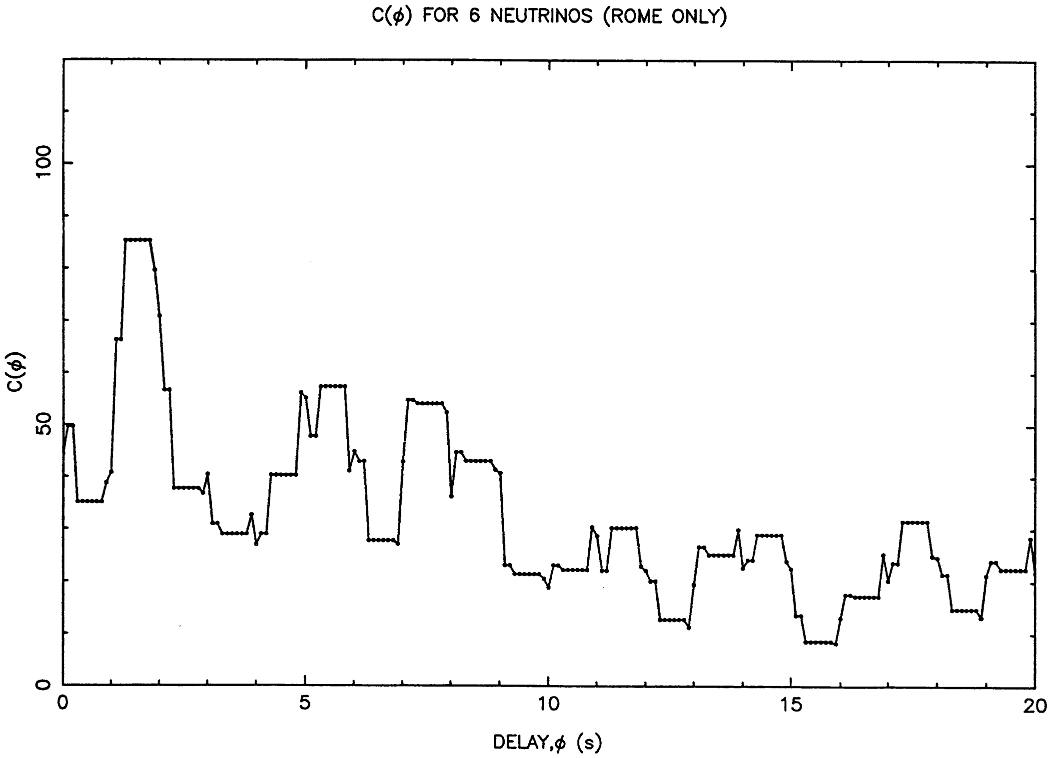
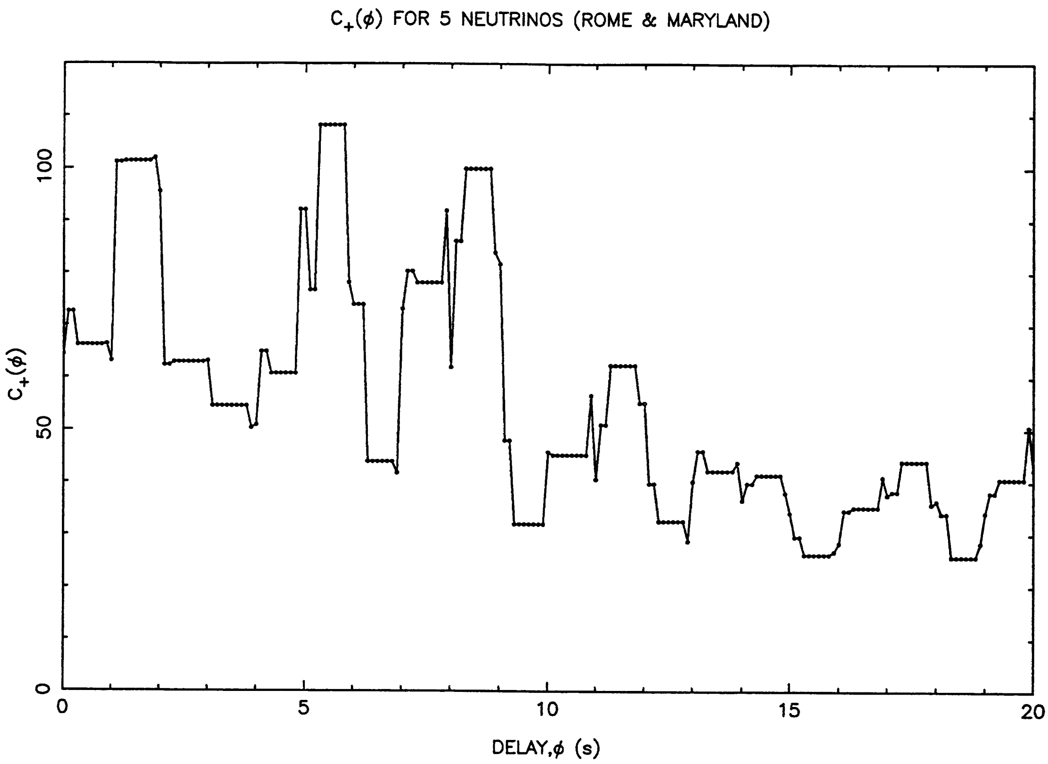


Figure 2.3: Searching for good time delays using the net-excitation method applied to the data set of Fig. 2.2, using Rome data only. (a) contains all six "neutrinos" seen in Fig. 2.2, while (b) omits the isolated "neutrino" event near 2h 52m 2s.



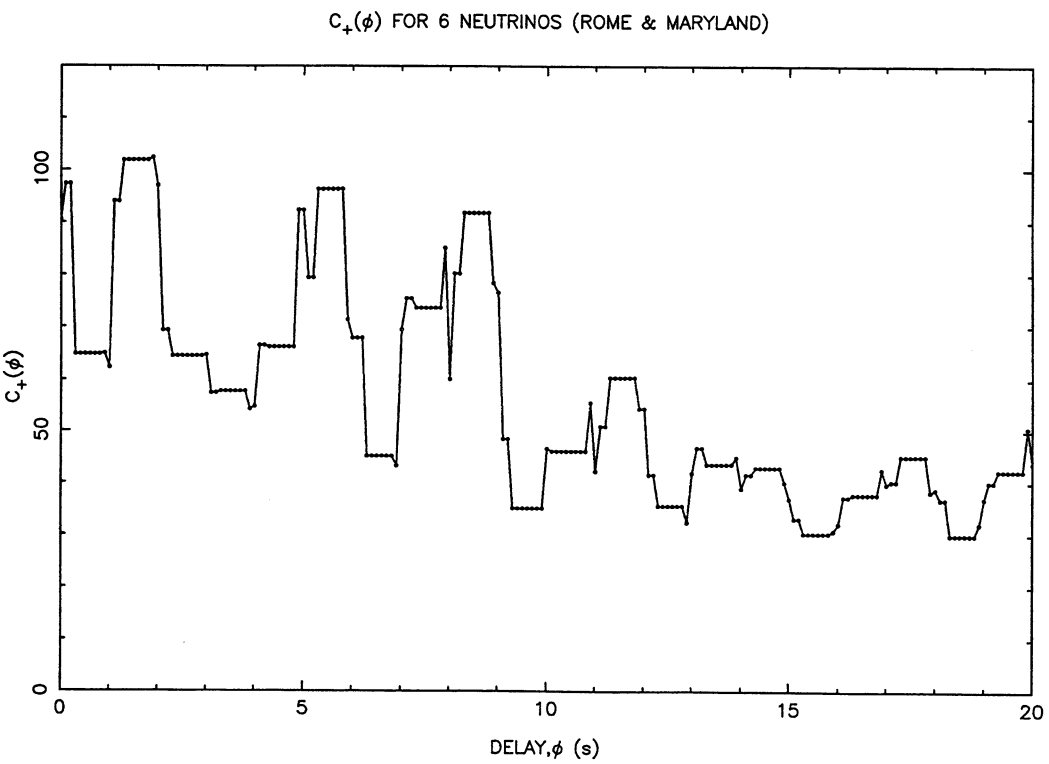


Figure 2.4: Searching for good time delays using the net-excitation method applied to the data set of Fig. 2.2, using both Rome and Maryland data. (a) contains all six "neutrinos" seen in Fig. 2.2, while (b) omits the isolated "neutrino" event near 2h 52m 2s.

near 1.5 s and those near 5.5 s and 8.5 s. In fact, if one uses only the 5 "burst neutrinos", the *best* time-delay is 5.5 s.

One could argue, therefore, that based on RTM's own selection criterion, they could have changed the delay time from 1.4 s to 5.5 s on receipt of the Maryland data. In fact, to us it seems natural to try to match the "double neutrino" event (arrival times 40.6 s and 41.0 s) with the highest Rome peak at 35.3 s, leading to a delay of about 5.5 s, which as we have seen is as good as or better than the delay they originally chose.

This is not, of course, to argue in favor of the reality of correlations at other delays. Our point here is to show that the range of time-delays that were open to RTM was considerable. Had, by accident of the noise in the gravitational wave detectors, the time-delay at 5.5 s proved a bit more significant, RTM would presumably have had no problem justifying its adoption. The physical model that they offered as a possible justification for the 1.4 s delay — that a small neutrino mass delays them relative to the gravitational waves — is untenable on other grounds (see our Section 2.1), and in any case it could surely have been stretched to justify a 5.5 s delay. Other ad hoc models, perhaps invoking unknown particles that excite the gravitational wave detectors, could easily have been devised to justify either advances or delays of small or moderate size.

We believe, therefore, not only that much larger values of $|\phi|$ could have been defended, but, indeed, that they *should* have been thoroughly examined by RTM once a time-delay model was adopted for analysis. As we have seen, RTM did indeed perform such an examination.

Our choice of delay. Consequently, we must regard the delay between gravitational waves and "neutrinos" as a free parameter like the threshold, and we choose the most favorable delay (within a pre-determined range) for a given set of random data.

We fix the range of available time-delays by staying with the RTM model (Amaldi *et al.* 1987) of ascribing the delay to the effect of a neutrino mass, m_ν. The time-delay between a gravitational wave traveling at the speed of light and such a neutrino with energy E_ν after traveling a distance d is

$$\delta t = \left(\frac{m_\nu c^2}{E_\nu}\right)^2 \frac{d}{2c}. \tag{2.9}$$

We need only fix an upper bound on the allowed mass and adopt a value for the typical energy of the neutrinos. By changing RTM's value of 10 eV for the maximum neutrino mass

(Amaldi *et al.* 1987) to a still reasonable 20 eV, and by relaxing the RTM energy estimate of 10 MeV to the actual measured average energy of the five Mt. Blanc burst events (8.4 MeV), we broaden RTM's allowed range of (0, 2.7) s to (0, 15.3) s. Hence we have run our main Monte Carlo experiment with the choice of delays

$$0.0 \text{ s} \leq \delta t \leq 14.9 \text{ s}, \tag{2.10}$$

in steps of 0.1 s.

For this parameter we feel we may have been conservative, i.e., that we could have defended wider ranges and hence obtained even larger corrected probabilities for the correlations. One could argue that negative delays (neutrinos preceding gravitational wave "events") should have been considered, since the new physics required to explain any correlations might well involve a new elementary particle that excites the gravitational wave antennas, and this might have traveled more slowly than the neutrinos. By the same argument, the time delay between neutrinos and the new particle could have been very much greater than the limits from the mass of the neutrino, since the new particle's mass could be very much larger. Without an *a priori* model for the physics of these correlations, it is hard to argue for any restriction on the time delay. Instead, a more practical reason for our accepting the relatively narrow range of 15 s is that RTM would probably not have looked for time delays at all had not the peaks in the gravitational wave stream been fairly near the "neutrinos" in Fig. 2.2.

We shall argue, shortly, that if one adopts a different range, the probability just scales in proportion. For example:

1. (conservative scenario) if one feels that delays in the range (−60 s, +60 s) are suitable, and that this range could reasonably have been searched, then the "true" probability will be larger by about a factor of 8 than the one we derive in Eq. 2.12 below;

2. (RTM scenario) if one feels that RTM's original eyeball estimate was binding, ±0.5s, and that during subsequent analysis no other delay could have been considered, then the "true" probability will be smaller by a factor of about 15 than the one we derive.

This illustrates how hard it is to estimate realistic probabilities when data have been analyzed by *a posteriori* criteria.

Note that in Section 2.5.1 we attempt to calculate the *a priori* probability of the correlations in the 2 hour data set, in a way which is independent of one's guess as to the available choice of time delay. We do this by removing those 50 s of data which RTM inspected to choose their delay of 1.4 s, and testing the predictive power of this delay on the rest of the two hours of data. We find the results are similar to those in our simulations that use a range of delays of 15 s.

Our algorithm

Having decided on the ranges of our free parameters, we proceeded as follows. Each simulated data set consisted of two gravitational wave streams and one "neutrino" stream generated as described in Section 2.4.2. For each threshold, we searched through the whole range of time-delays to find which one gave the best correlation as measured by the threshold-coincidence analysis method, and then we calculated the *apparent* probability of this correlation using Eq. 2.7. We performed the same analysis for each allowed threshold, and selected from all the one which gave the smallest apparent probability. We repeated this for each Monte Carlo data set (150 sets in our first run, 10^4 in our second) to see how often apparent probabilities smaller than any particular value occur. This allows us to correct the apparent probabilities for RTM's freedom to choose thresholds and time-delays, a freedom they did not systematically quantify. We assume that the *relative frequency* of any apparent probability in our simulation is the *true* probability that that sort of correlation will arise by chance in a given random set.

2.4.4 Results

We performed two simulation runs, the first using 150 data sets and the second with 10^4. We made minor changes between the two, primarily in the range of time-delays we accepted. Because one of the difficulties of understanding the significance of any statistical analysis is knowing what analyses have been performed and *not* reported (Section 2.5.1 below), we report *both* of our analyses here separately. We have not performed any others.

First simulation run

In our first run, we permitted the delay to vary from -60 s to $+60$ s in steps of 0.1 s. Although this range is larger than we have argued for, it is clear that, since each data point in the simulated time-series is independent, the coincidences found for different delays will be uncorrelated if the delays differ by more than 1 s. Therefore the probability of obtaining a given number of coincidences will simply scale linearly with the number of choices of delay. Searching 150 data sets over a range of 120 s is equivalent to searching 1200 data sets over a range of 15 s, which is the range we adopted for our second run. The first run therefore contains 12% as many independent trials as the second one. We regard one trial that uses a 15 s range of time delays and the range of threshold values described earlier as roughly equivalent to one RTM experiment.

We would therefore expect to find only correlations that have *true* probabilities of the order of 10^{-3} in our first simulation. In fact, we found one data set that had correlations that had an *apparent* "probability" that was even smaller than that of the RTM correlations!

In each of the 150 random data sets we summed the two gravitational wave streams and searched above the selected threshold for coincidences with neutrinos at the appropriate delay. The least probable correlation occurred in data set 55: at threshold 110 K and at delay 28.0 s, we found 22 gravitational wave–neutrino coincidences. In Fig. 2.5 we present these results in the same way as is done in Fig. 14 of Aglietta *et al.* (1989).

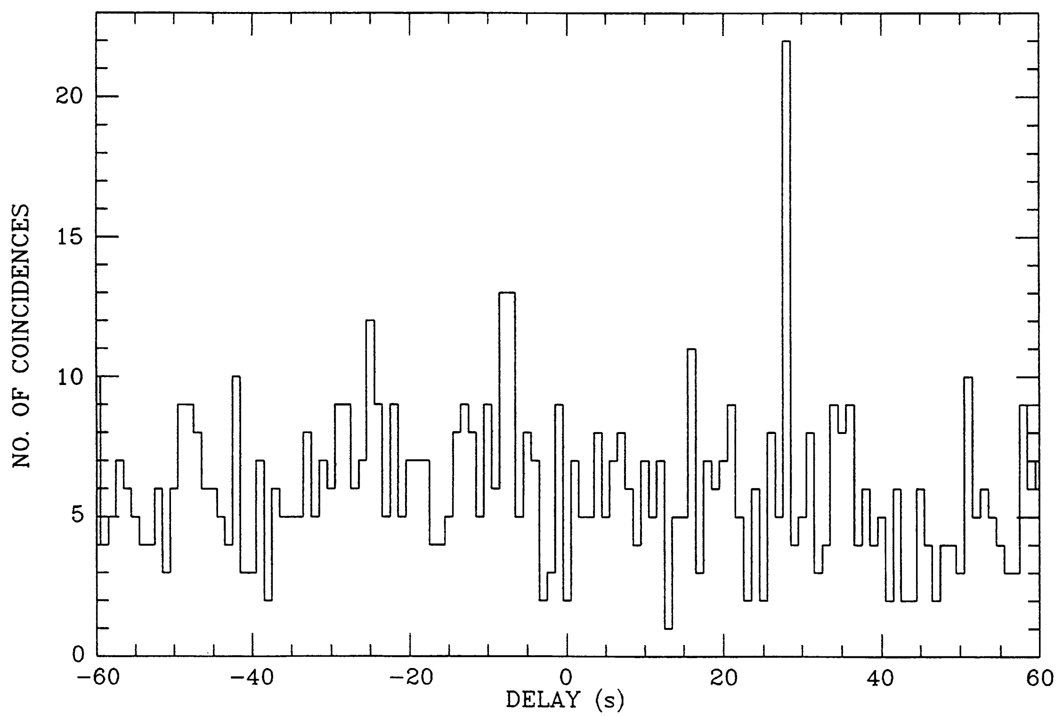


Figure 2.5: Histogram of number of threshold-coincidences against delay time for data set 55 of the first simulation run.

For this data set, there were 86 simulated "neutrinos" within the two hours, and at threshold $T = 110$K there were 512 "gravitational wave events", giving an expected number of coincidences $\bar{n} = 6.116$, by Eq. 2.6. The Poisson probability of obtaining 22 coincidences here is (from Eq. 2.7)

$$p_{\text{lowest}} = p_{\bar{n}=6.116}(22) = 5.3 \times 10^{-7}.$$

This is a *more significant* peak than that found by RTM, using RTM's method of calculating the probability, although we found it in the equivalent of only 1200 experiments.

We then submitted this data set to a net-excitation analysis, using the summation method and using the "best" time-delay of 28.0 s. A plot of our results in the style of Fig. 11 of Aglietta *et al.* (1989) appears in our Fig. 2.6(a). In Fig. 2.6(b), we reproduce the original RTM figure itself. There is a remarkable similarity between the two. The actual value obtained for $C_{+}(28.0)$ was 72.2 K, easily larger than any of the 10^6 background values with which it was compared to generate Fig. 2.6(a). We are confident that we could have made the trough in this figure even lower, had we generated more comparison values. We conclude that in roughly 1200 experiments, we have found correlations as strong as than those RTM found in the real data. Note that this was the *first* time we had performed a net excitation analysis, and the only time for these datasets. It is conceivable that there were other datasets in this experiment with net-excitation correlations this strong, and that the threshold coincidence method is an inefficient way of finding them.

However, it is not possible to draw reliable conclusions on the basis of one unusual data set, so we returned to the computer and did a longer simulation run.

Second simulation run

At the outset of this run we decided that the narrower range of time delays of 0.0 to 14.9 s would be more appropriate for simulating the RTM procedure. We performed 10^4 simulations in order to improve our statistics. We still found only one data set which was less probable than the real data, using the threshold coincidence method, but we found several with only slightly larger probability. These have enabled us to form a reliable estimate of the frequency of occurrence of these low-apparent-probability data sets.

The most improbable simulated data set. The most improbable data set in our second run was number 327, which had a peak of 9 coincidences at delay 7.6 s at a threshold

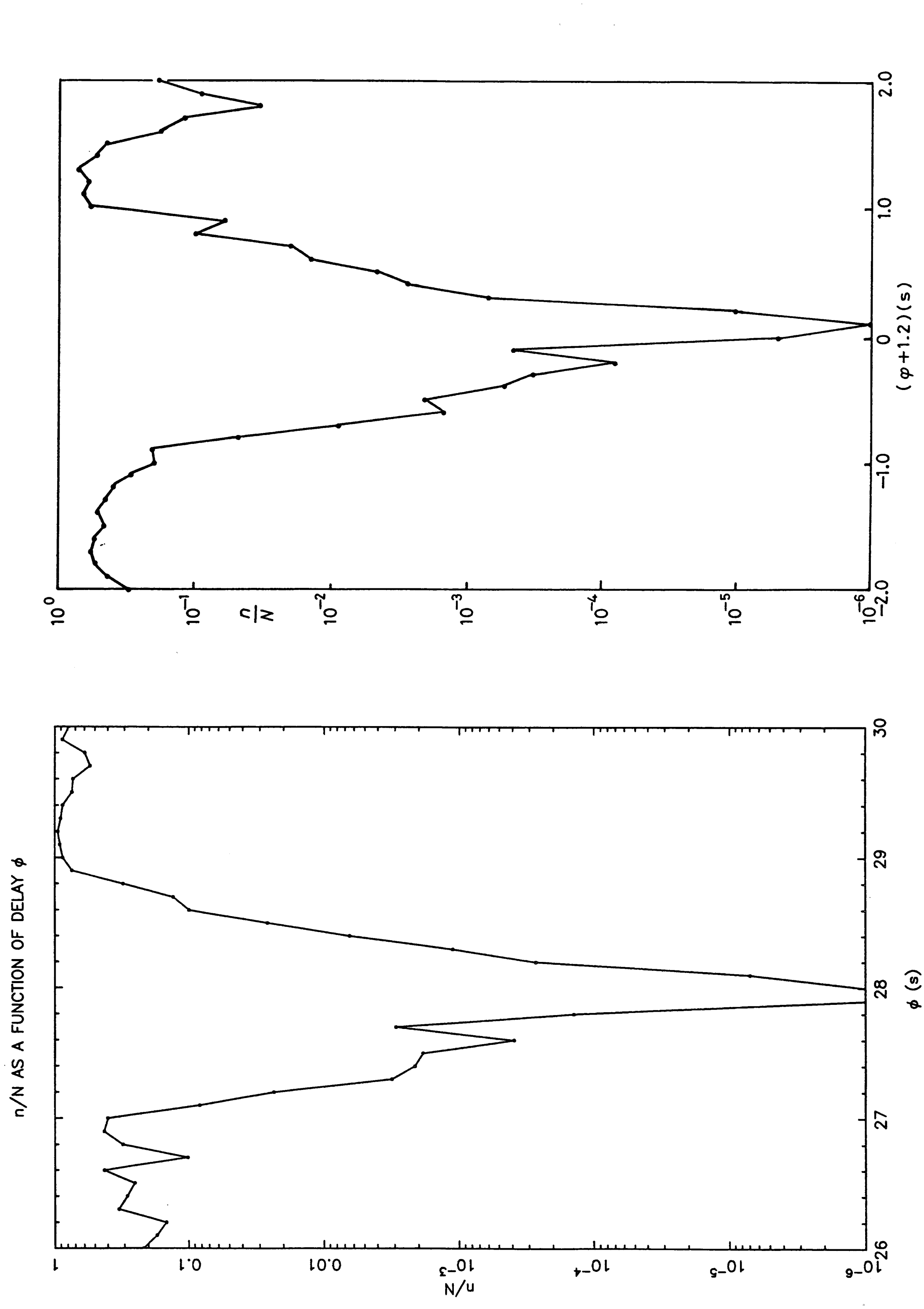


Figure 2.6: Result of the net-excitation analysis of simulation set 55 (a) compared to the RTM analysis of the real "neutrino" and gravitational wave data (b). ((b) reproduced from Ref. Aglietta *et al.* 1989 with permission.)

170 K. The histogram of coincidences against time delays is in Fig. 2.7(a), plotted with the corresponding one for the RTM data (b). There were only 77 neutrinos in the two hours of simulated data, and 83 gravitational wave events above this threshold. The number of expected coincidences is $\bar{n} = 0.888$ [Eq. 2.6]. From Eq. 2.7, the peak of 9 has a probability of 4.5×10^{-7}, less than that of RTM's correlation.

When we applied the net-excitation analysis to this data set, the result was quite different from that for our earlier data set: the dip in Fig. 2.8 is by no means as dramatic as it was for the RTM data, or for our own Fig. 2.6(a). Although there are an unusual number of coincidences in this data set, the average excitation of the gravitational wave detectors was not extraordinarily high at the (delayed) time of neutrino arrivals. This illustrates simply the fact that the two analysis methods measure different, albeit related, properties of a data set, and so simple probability estimates based on one or another of these statistics will not necessarily agree.

The relative frequency of occurrence of such correlations. Given the pseudo-random neutrino and gravitational wave data sets, each threshold T on the gravitational wave data stream determines an expected number of coincidences $\bar{n}(T)$. Choosing a delay ϕ then fixes the actual number of coincidences $n(T, \phi)$. We seek the lowest apparent Poisson probability over all thresholds and delays, which we call q:

$$q = \min_{\text{thresholds } T} \left\{ \min_{\text{time-delays } \phi} p_{\bar{n}(T)}[n(T, \phi)] \right\}, \tag{2.11}$$

where $p_{\bar{n}}(n)$ is given by Eq. 2.7. The frequency distribution of values of q in the 10^4 data sets gives us our realistic probability distribution. One would expect this to be proportional to q, if the RTM raw probabilities were realistic, so that smallest values of q occurred the least frequently. As Fig. 2.9 shows, the actual distribution of q is just the opposite: the freedom to adjust parameters makes small values of q very much more probable than large ones.

The analytic form of this distribution is not known, but Fig. 2.10 shows that for most of the range of q the curve is fairly close to being exponential.

Our interest is in the smallest values of q, whose histogram is plotted in Fig. 2.11. Within the statistical fluctuations, the distribution is fairly flat, which is what we would expect if the behavior as $q \to 0$ is a regular extrapolation to zero of the low-q trend in Fig. 2.10, and does not become singular as $q \to 0$.

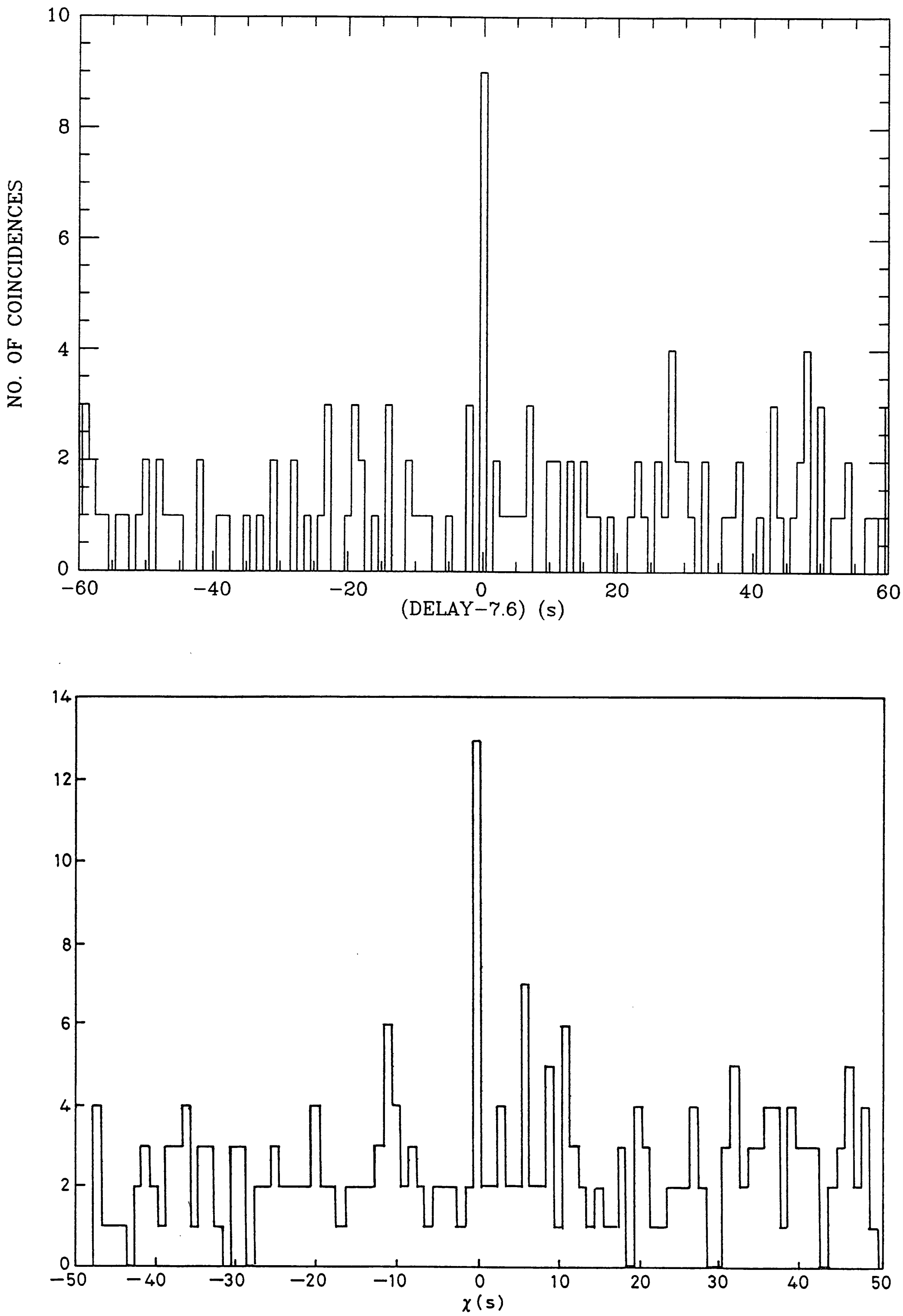


Figure 2.7: Delay histogram for data set 327 of the second simulation (a) compared to the RTM histogram of the real data (b). ((b) reproduced from Aglietta *et al.* 1989 with permission.)

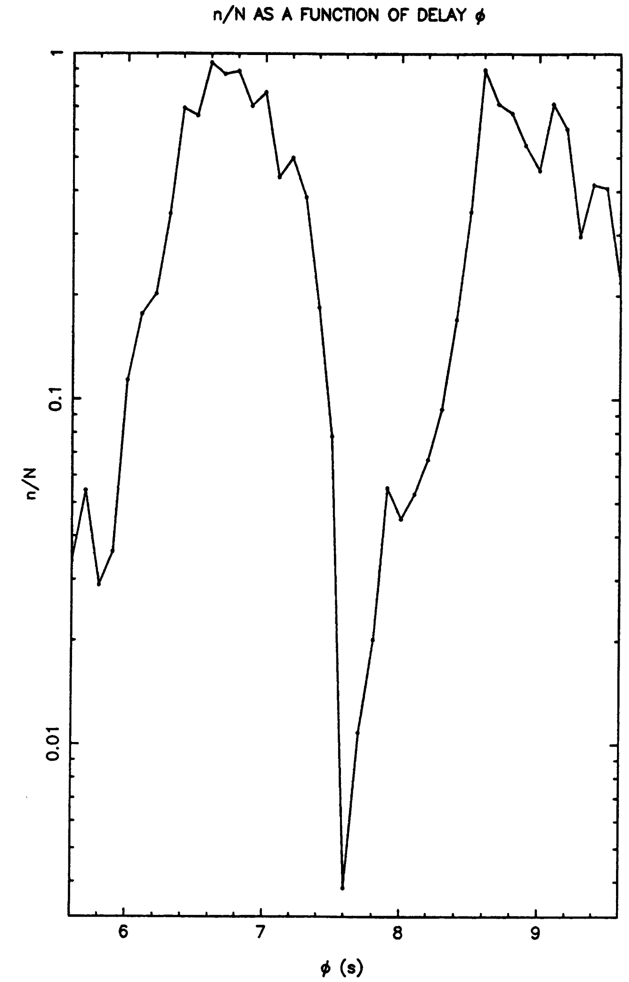


Figure 2.8: Net-excitation analysis of set 327. Although the threshold-correlation method gives as strong a correlation here as for set 55, the net excitation analysis does not show nearly as dramatic a dip as in Fig. 2.6

We can use this figure to estimate the realistic chance probability of the threshold coincidence correlations in the RTM data as follows. The first 20 bins in Fig. 2.11 contain 21 data sets. This suggests that the realistic probability is that any one bin will contain one data set in each 10^4 trials. Since the width of each bin is $\Delta q = 10^{-6}$, the true probability that a data set will give q less than 10^{-6} (i.e., will fall in the first bin) is

$$p\left(q < 10^{-6}\right) \approx 10^{-4}. \tag{2.12}$$

Thus, a more realistic estimate of the a priori *probability that the two hours of data which RTM analyze will show the sort of threshold coincidence correlation they find is* 10^{-4}. This estimate does not, of course, allow for other effects, such as the selection of the data set and

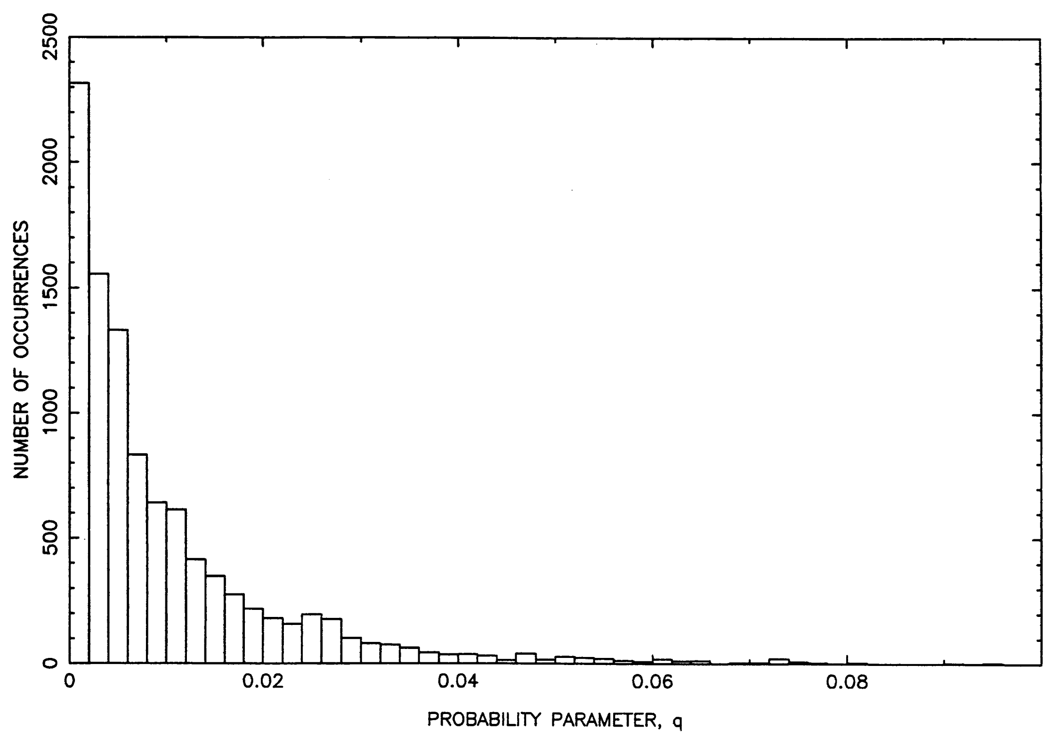


Figure 2.9: Relative frequency distribution of the values of the parameter q in our second simulation run. This parameter is used by RTM as their probability estimate. If this were the true probability, this figure would be a straight line through the origin.

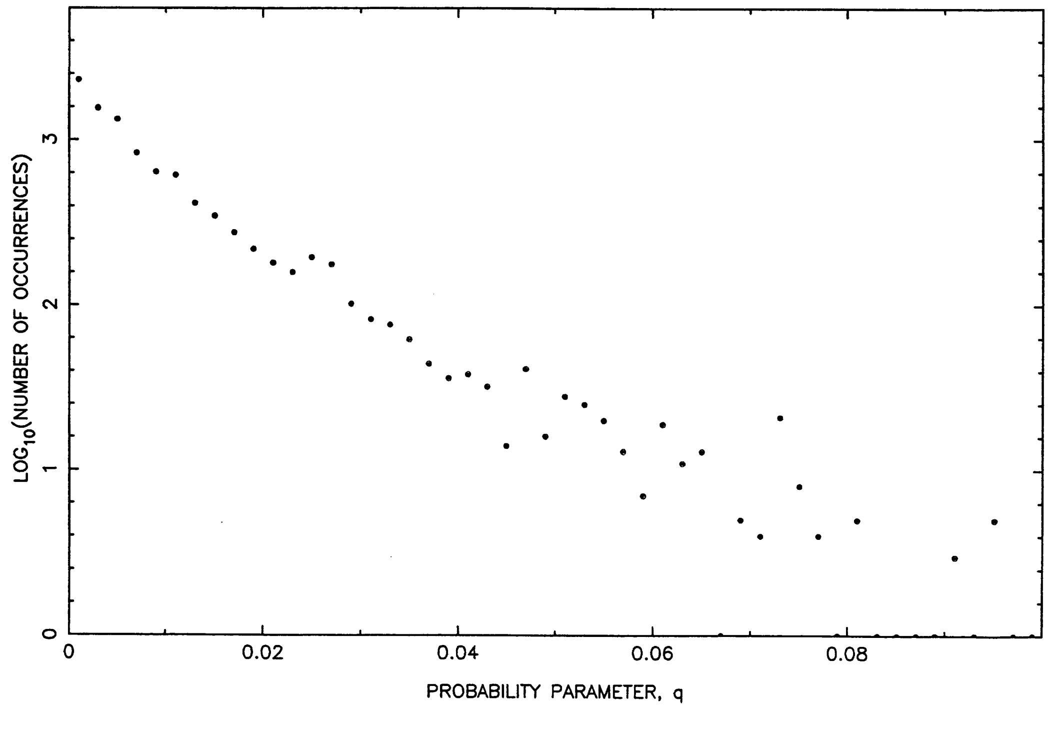


Figure 2.10: Logarithm of the previous figure, showing a nearly exponential distribution.

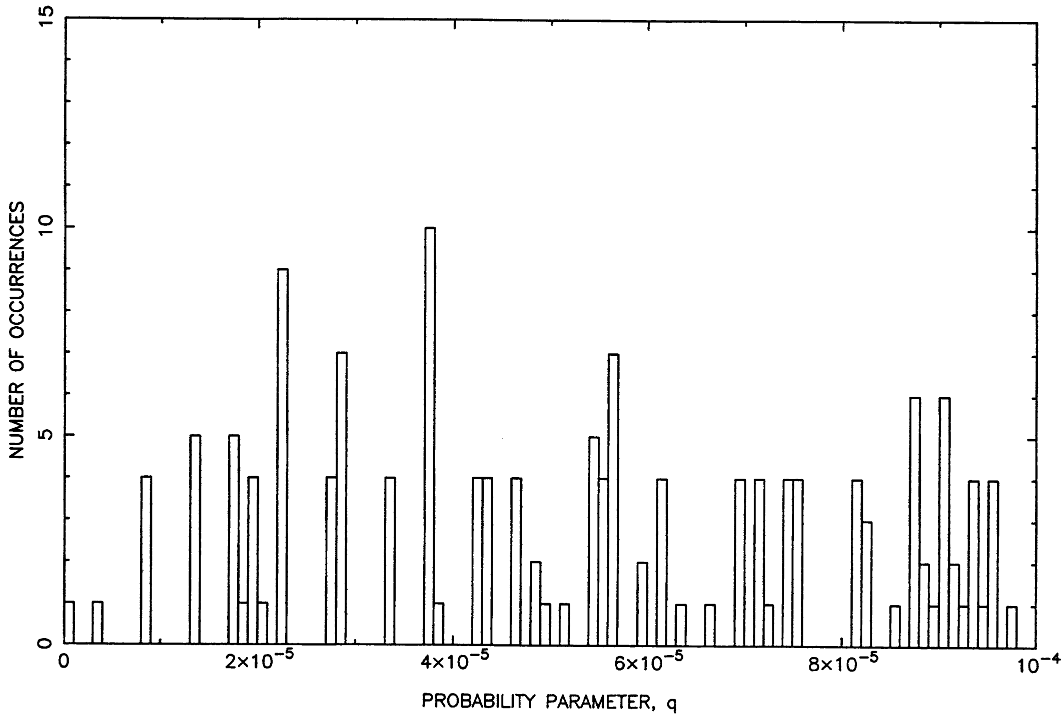


Figure 2.11: Frequency distribution of q for small q, allowing an estimate of the distribution of unlikely correlations. If the distribution in the previous figure is fit by a straight line for small q, then its slope in this figure would be nearly horizontal because of the greatly enlarged scale for q.

the *a posteriori* nature of the analysis method. In the next section we take these less easily quantified effects into account. We shall also show that it is possible to find evidence within the RTM threshold-coincidence analysis itself that our simulation probabilities are closer to the true probabilities than RTM's own estimates.

2.5 Reassessment of RTM correlations

We shall now make a full reassessment of the probabilities of the correlations RTM have found, in the light of our Monte Carlo simulations. We shall study the results of five papers, all of which reported coincident events: two (Amaldi *et al.* 1987; Aglietta *et al.* 1989) found coincidences between the gravitational wave detectors and the Mt. Blanc neutrino detector (see Section 2.5.1); two (Amaldi *et al.* 1989; Aglietta *et al.* 1991b) found coincidences between the gravitational wave detectors and the Kamiokande particle detector (see

Section 2.5.2); and one (Aglietta *et al.* 1991a) found coincidences between gravitational waves and IMB (see Section 2.5.3). (Another paper (Amaldi *et al.* 1988) found correlations between the two gravitational wave detectors themselves, but the probabilities found were not so unusual, so we review it briefly in the Appendix 2.7.2.)

We shall deal in this section with the main analysis methods which RTM use; though in the interests of completeness, we have included many of the details of the various analysis papers in the appendix at the end of the paper. We shall first reassess the Mt. Blanc neutrino–gravitational wave coincidences; then we shall re-examine the Kamiokande– and IMB–gravitational wave coincidences.

2.5.1 Reassessment of Mt. Blanc–gravitational wave coincidences

The RTM calculations of probability (Aglietta *et al.* 1989) are seriously affected by certain *a posteriori* choices they have made. Using our simulations in Section 2.4, we have already assessed the effects of some of these choices—delay time and thresholds—on the results of their threshold-coincidence analysis, coming to the conclusion that the correlations they find have an *a priori* probability of about 10^{-4} in any single random data set. We have also shown that probabilities derived from the net-excitation method are not reliable below values of a few times 10^{-4}.

In this section, we firstly examine the behaviour of RTM's data set without the 50 s of data which they used to choose their first time delay of 1.4 s, and we attempt to use this data set to test the predictive power of their choice. This gives an estimate of the probability of the RTM correlations which is independent of one's guess as to how much freedom RTM had to adjust their delay parameter.

We then consider other significant *a posteriori* choices that RTM made that made it easier for them to obtain correlations. We shall see that some are quantifiable, while the effects of others can only be guessed at. The overall effect of these considerations is further to increase the likelihood of RTM's discovering the correlations which they find.

Contamination of 2-hour data by including the "eyeballed" set.

An important issue is the fact that RTM included in their full data sets the original 50 s stretch of data that contained the 5ν burst that originally suggested to them that they should search for a time-delay of about 1.4 s. RTM are aware that this biases their probabilities and at one point attempt to show that this has a negligible effect on the final result. We will explain below why their argument is wrong. We will then show how removal of the 50 s of data can be used to control for RTM's ability to choose the time delay, by assessing the predictive power of a 1.4 s delay chosen from those 50 s of data, used for the whole data set excluding those 50 s.

Effect of contamination. It is straightforward to estimate the effect of this contamination on the threshold coincidence method that RTM apply to the Mt. Blanc data. In the 2-hour stretch they analyze, they find 13 coincidences at the adopted threshold. Against an expected value of 2.29, Eq. 2.7 gives a chance probability of about 10^{-6}. If we exclude the first "neutrino" of the Mt. Blanc burst, which is clearly in coincidence with gravitational waves in summation above the threshold of 150 K (and is the only one), then the number of coincidences at this threshold falls to 12. This gives a chance probability of about 5×10^{-6}. This is before corrections for the arbitrariness of the threshold, time-delay, etc. The contamination thus makes their threshold-coincidence probabilities a full factor of 5 too small.

The contamination is much greater in the net-excitation method. Consider the statistic $\sum_{i=1}^{N_\nu} E(t_i + \phi)$ in Eq. 2.1 which seems to give such an unusually large value. The five "neutrinos" originally "eyeballed" in Fig. 2.2, with the delay deliberately chosen so that the Rome gravitational waves so delayed with respect to the "neutrinos" are appreciably higher than the average background, will each add about $(82.4 - 28.6 \approx)$ 55 K extra to this sum (see Fig. 2.4) at delays of both 1.4 s and 1.1 s. This artificially increases the sum by about 275 K and so, when divided by 96 for the number of "neutrinos" detected in the two hours under analysis, this contributes about 3 K to $C_+(1.1)$ and $C_+(1.4)$ (see equation Eq. 2.2). This would considerably alter the ranking order of $C(1.1)$. Fig. 12 of Aglietta *et al.* (1989) shows that if the value of $C(1.1)$ were 3 K less, there would be about 20 "background" values greater than $C(1.1)$, while there were none before. That is, without the "eyeballed" data, the RTM estimate of the probability of the correlations in the rest of the 2-hour data would be raised to 2×10^{-5}. This factor of 20 still leaves the probability below the range of reliability of the net-excitation method.

Contamination correction as a way of controlling for the time-delay freedom. Excluding the "5-neutrino burst" is in fact another way of compensating for the freedom to choose the time delay in the correlation analyses. If we allow only the original RTM "eyeballed" time-delay of 1.4 s and exclude that data set from the subsequent analysis, we would obtain an unbiased result that tests the ability of the original set to predict correlations in the extended data set. For the net-excitation method, this would remove the principal degree of freedom. However, all we have been able to do is perform that test for the revised delay of 1.1 s, where we found the probability went up by a factor of 20. We should really apply this correction to the original time-delay of 1.4 s, but the RTM papers do not provide enough information for us to be able to do this. However, we can be certain that the proper correction would raise the probability even further, since in the full data set a delay of 1.1 s gave a better correlation than did 1.4 s; while in the data that one removes (containing the 5 ν burst), the 1.4 s time-delay was better (see Section 2.4.3).

The threshold-coincidence method is, of course, also contaminated by this, and we have seen that this correction is a factor of 5. This is a correction only for the freedom to choose time-delays, not for the threshold freedom. Since in our simulations (which are not affected by this contamination because we do not look at the first 50 s to get a time-delay and then re-use this stretch of data in estimating the probability that the full set shows a correlation) we took a correction factor of 15 for time-delays (a 15 s span rather than RTM's 1 s), the factor of 5 takes the corrected RTM probability most of the way toward our simulation estimate. Moreover, the remarks in the last paragraph about using the original time-delay of 1.4 s apply here too. This will raise the correction still closer to (if not beyond) our factor of 15. *We find, therefore, that the contamination effect can be used to control for the time-delay freedom, and when one does so one finds consistency with the probabilities of* 10^{-4} *produced by our simulations.*

Problems with the RTM contamination correction. RTM realized that the contamination of the 2-hour data by the "eyeballed" data was a problem, and they attempt to show that it does not really change things by calculating the net-excitation ranking statistic (comparison of $C_+(1.1)$ with random "background" values) with and without the 5 "neutrinos" of the Mt. Blanc burst. They find no significant change (Fig. 5 of Aglietta *et al.* 1989). However, the comparison is flawed because they used only $N = 10^3$ background values to calculate the "probability" of the correlation, both with and without the 5ν burst. Such a calculation can (according to our argument on the independence of the background values) indeed distinguish between data sets that have a chance probability greater than about 10^{-3}, but unfortunately RTM adduce this calculation as evidence that a data set with a probability of 10^{-6} is uncontaminated. Even if their method were reliable, they would have had to have used at least 10^6 background points to have drawn any conclusions.

RTM tell us (Pizzella & Pallottino 1991) that they have, in response to our criticism, subsequently performed such an analysis with 10^6 points and find that their net-excitation probability goes up by a factor of 5 when the original "neutrinos" are excluded. While this takes them some of the way toward the 10^{-4} level that we feel the correlations really warrant, they still have not compensated for changing from 1.4 s to 1.1 s, and they are, in any case, using a method whose probabilities are unreliable at this level.

Further corrections to the probability of the Mt. Blanc correlations

We have shown from our simulations that the correlations in the Mt. Blanc data occur with probability $\sim 10^{-4}$. We have confirmed this by removing the data from which the 1.4 s delay was chosen, and testing the predictive power of this delay on the rest of the data. We start this section, therefore, with the estimate that, for the given Rome-Maryland-Mt. Blanc data set, the probability that RTM would have found the correlations they did find is about 10^{-4}.

Selection of the data set to analyze. Through our simulations and our attempts to correct for time-delay and threshold freedom in the RTM analyses, we have arrived at the conclusion that the given Rome-Maryland-Mt. Blanc data set contains correlations with a real probability of about 10^{-4}. While larger than the RTM claim of 10^{-6}, this is still potentially significant. However, we now have to turn to a number of corrections that have to do with other *a posteriori* choices made by RTM.

The first is that RTM see their correlations only in a particular 2-hour stretch of data, which was not selected because of any property of the 50 s "eyeballed" data set. Indeed, RTM looked for correlations in other, earlier, data sets and found none at the same time delay of about 1.2 s. Also, they examine longer and shorter data sets and find that the effect becomes much weaker for periods less than about 50 minutes and greater than about 150 minutes (Fig. 9 of Aglietta *et al.* 1989). Indeed, they seem to regard this as evidence for the reality of their correlations, since if they were associated with the supernova, then one would expect them to be transient.

However, when assessing the significance of correlations, one must be careful to start from the *null hypothesis*, that the correlations arise by chance. Then it is clear that one's ability to choose the data set in which one finds correlations is another free parameter, like the time-delay itself. Since one has no *a priori* idea of the length of the period during which these correlated "neutrinos" and gravitational waves (or new particles) should have been emitted by the supernova, it is fair to expect that if correlations as strong as the ones RTM found had appeared instead in, say, a longer or shorter stretch of data, RTM would have treated them just as seriously.

In fact there are *two* variable parameters here: the length of the data set and its starting time. RTM make a natural *a priori* choice in selecting a data set which includes the Mont Blanc 5ν burst, but it need not have been 2 hours long and it need not have been *centered* on the burst. It would have also been natural to have looked for phenomena either immediately preceding the putative collapse event or immediately following it; indeed, on physical grounds it seems rather unlikely that any correlated phenomena would have occurred *both* before and after the collapse, since the physical conditions are so different on either side.

The length of the data set is even more important. RTM analyze a 2 hour data set, but again give no physical reason for having made this choice *a priori.* The reason for this choice seems to come from Fig. 9 of Aglietta *et al.* (1989), where, for a fixed delay of 1.2 s (given by "eyeballing"), RTM compare $C_+(\phi)$ and $C_+(\delta_1, \delta_2)$ for different values of the length of the data set, and show that the best correlations for the net-excitation algorithm occur for lengths between about 100 and 130 minutes, with the "probability" of the correlation increasing fairly sharply by about 1 or 2 orders of magnitude outside a window from about 70 to 150 minutes. RTM apparently used this information to select the data set they analyzed. In fact, RTM stress that the 2 hour length of data is not optimal: 135 minutes is better. But it is clear that even so they have made a considerable optimization by choosing a value near the "best" one, when they could have chosen a length of anything from, say, a few minutes to 36 hours.

We have not attempted to simulate this freedom to choose the data set in our Monte Carlo analysis; it would have been computationally very expensive. We also do not know from the published papers how many data sets RTM actually looked at. In the absence of simulations, the following argument gives us some idea of the size of the effect.

We would like to know how many essentially independent data sets RTM could have analyzed. Let each data set contain the Mt. Blanc burst, and let us take a minimum reasonable data set length L to be 8–10 min. If we enlarge the set by a factor of 2, the larger set will have statistics reasonably independent of those of the smaller included in it. Each such doubling of the length produces a new "independent" set, until L reaches 36 hours, the total of the data apparently available to RTM initially. This requires 8 doublings, giving 9 sets. For the shorter sets there are actually two independent sets, one ending with the Mt. Blanc burst and the other beginning with it. Doubling these "post-Mt. Blanc" sets until the Maryland detector goes off line because of its electrical problems adds 5 more sets, giving 14 in all.

We shall therefore take a factor of 10 to be a reasonable lower limit on the correction we need to make for this selection effect. *This raises our estimate of the probability that RTM's analysis methods would have found correlations in entirely random data to about* 10^{-3}. Next we turn to the problem that their analysis methods were themselves invented *a posteriori.*

"Trial and error" analysis. Every textbook introduction to statistical analysis emphasizes the problem that, the more often one analyses a given set of random data in different ways, the more likely it is that one will uncover a correlation of apparent significance. In our simulations in Section 2.4 we have accordingly reported *all* the trials we did. Unfortunately, it is impossible from RTM's papers to learn whether they performed other analyses of the data that they do not report. We have indicated at several places in this paper our guess that they may (or even should) have done so.

For example, the most natural kind of analysis to have done with two gravitational wave streams and the Mt. Blanc data is a triple-coincidence analysis, where one identifies gravitational wave "events" by setting a threshold separately on these two data streams. The threshold need not be arbitrary: a reasonable one is a level where one expects only a few coincidences over the selected data set if the data are random (a low "false alarm rate"). RTM do not report such an analysis. Instead they report a double coincidence analysis in which the gravitational wave data are added together before being thresholded, and they search many thresholds. They also introduce a non-standard method, the net-excitation method. However, as we shall see, RTM *do* report having done such a triple-coincidence analysis for the Kamiokande data and gravitational wave detectors.

Having used certain methods for the Mt. Blanc data, they then do not stay exclusively with them for the KII data. The net-excitation analysis is done but not examined in detail. The threshold-coincidence method is not reported, but the results of the triple coincidence method are. And the length of the data set is changed. The papers do not tell us if RTM performed, say, the threshold-coincidence analysis of the KII data over the original time-span and did not report it because the results were not very significant.

One cannot argue that these tests are all roughly equivalent, so that if a correlation shows up in it will show up in all: this is not necessarily the case. For example, our Monte Carlo simulations produced two "good" correlations as measured by the threshold-coincidence method, but one of them gave a good correlation using the net-excitation method and the other did not. Here, the choice of the analysis technique used makes a difference of a factor of 10^3 in the "probability" obtained. These methods all measure different things (though some methods are partly dependent on each other in ways which are not clear). So the significance of a reported correlation is diminished if other tests were applied that gave null or insignificant results, simply because the other tests *could* have given correlations (even if they did not).

Another worrying aspect of this is that there are occasions where it appears that a secondary analysis was designed after a primary analysis, and may therefore have been guided by the results. An example of this, which we have already seen, occurs in the design of the net-excitation method in Aglietta *et al.* (1989). When calculating the "background" neutrino–gravitational excitation to compare with the measured value given by Eq. 2.1, RTM make an unexpected choice: they use Eq. 2.2, in which the gravitational wave data streams are taken at different times, rather than simply shifting both gravitational wave data streams by the same amounts.

While this would not be unreasonable as an *a priori* choice (provided they had used enough data to ensure independence of the background values), the problem is that RTM by this time appear already to have performed the Maryland-Rome correlation analysis (Amaldi *et al.* 1988), which showed that the two gravitational wave detectors had an unusually high number of coincidences, at zero relative time delay, during the period under analysis. RTM should have known that by calculating the background as they have, they have obtained a marginally lower value for the apparent probability than if they had kept the two gravitational wave data streams tied together.

These points illustrate the difficulties that a posteriori *invention of data analysis methods can create. We find these effects impossible to quantify, but inevitably they raise the probability of finding correlations.*

Overall assessment of the probability

We now assemble our various corrections to arrive at an estimate of the chance that RTM would have found correlations of the level of significance that they have reported in entirely random data.

In the introduction to Section 2.5.1 we have put the chance probability of finding RTM-style correlations in a given random data set at about 10^{-4}. In Section 2.5.1 we have raised this to 10^{-3} because of their freedom to choose the data set. We cannot quantify the correction for what we have called "trial and error analysis", but it could be significant. One could conceivably get even another factor of 100 from this, since we actually found a variation of a factor of 1000 between different RTM analysis methods applied to the same simulated data set (see Section 2.4.4).

We therefore conclude that, because of the considerable freedom that RTM had (and frequently exercised) in looking for correlations, the a priori *chance probability of the correlations that they did find between the gravitational wave data and the Mt. Blanc "neutrinos" lies somewhere between 0.001 and 0.1.*

2.5.2 Reassessment of KII–gravitational wave coincidences

We turn now to the RTM coincidences between the Kamiokande detector and Mt. Blanc and the gravitational wave detectors. Of course if the correlations between the Mt. Blanc detector and the two gravitational wave antennae are due, even in part, to real neutrinos or other particles emanating from an astrophysical source, then similar correlations should be present in the data of other particle detectors, even though they did not exhibit obvious bursts of activity at this time, as did Mt. Blanc. RTM recognize that the correlations between the Mt. Blanc neutrino data and the gravitational wave data may have arisen by chance, so they rightly regard the acid test of the correlations to be their predictive power: do the particle data from Kamiokande show the same correlations with the same time delay? One would expect that the same methods of analysis as were previously used should yield correlations at a similar level of significance.

Unfortunately, RTM do not make a clean test of the predictions of the Mt. Blanc analysis: they do not simply apply the same analysis methods to the Kamiokande data as they used for the Mt. Blanc data. However, RTM do indeed claim to find correlations, with probabilities that lie between 10^{-3} and 10^{-8}, depending on the tests RTM apply. We shall reassess each of their methods in turn, but first we consider corrections that apply to all of them.

General corrections to the probabilities that RTM assign to KII.

Clock correction. As we mentioned before, the Kamiokande clock had an *absolute* timing uncertainty (difference from Universal Time) of ± 1 min, although *relative* timings were very accurate. This could be rectified to a certain extent by demanding that the Kamiokande neutrino burst at 7h 35m was coincident with the burst in the IMB detector, whose clock was accurate. RTM try to find a clock shift that will bring the two bursts of neutrinos detected at 7h 35m 41s (± 5 ms) and 7h 35m 35s ($\pm$ 60 s) UT in IMB and Kamiokande respectively into coincidence. But what does one mean by coincidence? Since the length of the Kamiokande burst (Hirata *et al.* 1987) was 13 s and the length of the IMB burst (Bionta *et al.* 1987) was 6 s, we have, unfortunately, a great deal of freedom in choosing when the two bursts should coincide. RTM use the criterion that the first neutrino of each burst should coincide, with a two second uncertainty; and hence that a reasonable range for the clock offset ϕ_c is the interval (5.7, 9.7) s. But they might have been equally justified in supposing that the *centers* of the two bursts should coincide, with an uncertainty of, perhaps, three seconds. This would have given RTM another six second window to search for coincidences.

Length of the data set. One of the most puzzling aspects of the RTM analysis of the KII data is that they do not stay with the 2-hour data set, but again give themselves freedom to choose its length: the one hour of data from 2h to 3h UT. There is no *a priori* reason for this, (all the particle data for the days 22 and 23 February 1987 were supplied by the Kamiokande group (Amaldi *et al.* 1989)) and no explanation is given, so we can only conclude that they have done it because it gives better results in the KII analysis. Indeed, in the one analysis where RTM repeat their analysis for the full 2 hours—the search for triple coincidences—their "probability" goes up by one or two orders of magnitude when they revert to the period they analysed in Aglietta *et al.* 1989. We can therefore expect that similar corrections of one or two orders of magnitude will apply to all "probabilities" quoted in the paper.

Change of analysis methods RTM use analysis methods for the Kamiokande-gravitational wave coincidences that are *different* from those that they used for the Mt. Blanc-gravitational wave coincidences. In fact they use several analysis methods. They do include the same "net-excitation" test, but for reasons that they do not explain, they do not give a detailed study of the probability of this correlation along the lines of their Mt. Blanc study. They also do not present a threshold coincidence analysis of the Kamiokande-gravitational wave data along the lines of the threshold analysis of the Mt. Blanc data, i.e., by putting a threshold on the summation or product data sets and looking for coincidences with the Kamiokande particles. This would seem to have been the most natural thing to have done. Instead they present a triple threshold-coincidence analysis.

Again, in the absence of any explanation, we must assume that RTM have done this to improve the results of their analysis. We shall try to correct for this where appropriate, and where possible, we shall use RTM's original analysis methods on the data they present.

KII-gravitational wave coincidences independent of Mt. Blanc

Net-excitation analysis. RTM performed the same net-excitation analysis on the Kamiokande–Rome–Maryland data as they had previously on the Mt. Blanc–Rome–Maryland data. However, RTM do not carry this through to a probability estimate. This is strange because, as we have said, such an analysis, *independent* of the Mt. Blanc data, is absolutely crucial in deciding whether this correlation effect is present in the Kamiokande independent of Mt. Blanc. Fortunately, RTM publish enough data (Amaldi *et al.* 1989) to allow us to do the appropriate analysis here.

In Fig. 1 of Amaldi *et al.* 1989, each value of $C_+(\phi)$ is compared with only 10^3 background values, and we see that at the chosen clock offset $\phi_c = 7.7$ s there is actually one background value which is larger than C_+. By RTM's method, Eq. 2.4, the probability of this correlation is about 2×10^{-3}. However, as we have seen, they should not put too much weight on the probability of a single point, when neighboring points separated by 0.1 s give different probabilities. In fact *all* the other points within a window ± 0.5 s of 7.7 s have associated "probabilities" of 10^{-2} or more, with several at 2 or 3×10^{-2} and one at 10^{-1}. The value at exactly 7.7 s cannot be physically significant, and a probability of about 2×10^{-2} would be more representative.

RTM do correct for the number of available choices of the clock time. Because they took

41 values of the clock time in the four second window about their chosen delay, separated by 0.1 s, they then multiplied all their derived probabilities by 41 to correct for this. Applied to their figure of 2×10^{-3}, this gives a probability of about 0.08. We must not apply the same correction to our best estimate of the probability, since we accept only 4 independent values of the delay. Multiplying our first probability, 0.02, by the number of choices, 4, also gives 0.08. This is a reasonable estimate of the significance of the Mt. Blanc-type correlations that exist in the 1-hour set of KII data selected by RTM. If in addition we allow for a further "washing out" of the correlations if we take the full 2-hour set (which could give up to a factor of 10 or so as outlined above), and if we include another 6 independent choices of clock setting, we see that *there is no evidence whatsoever for physically significant Mt. Blanc-type correlations in the Kamiokande data.*

We would be justified in stopping here and looking no further at coincidences between the Kamiokande data and either the gravitational wave data or that from Mt. Blanc. The Mt. Blanc correlations fail the "acid test" for their physical reality by being absent from the KII data. However, for the sake of completeness we shall continue briefly to review the other RTM analyses of the KII data.

KII–Maryland–Rome triple coincidences. Instead of doing the same sort of threshold–coincidence analysis that they did for the Mt. Blanc–gravitational wave data, RTM instead perform a straight triple-coincidence analysis, looking for threshold-crossings in both the gravitational wave data streams at times given by Kamioka particle times (with the appropriate delay). Indeed, this is the sort of analysis that we wish they had performed on the Mt. Blanc data in the first place. However, the fact that they did something different for Mt. Blanc makes it hard to offer this analysis as evidence that the Mt. Blanc correlations appear in Kamiokande.

RTM find a peak of 15 triple coincidences at a threshold of 40 K, for which they made an erroneous estimate of the probability, 8.2×10^{-6}, by using the mean rate of events appropriate to a different data set (see Amaldi *et al.* 1989 or our Appendix). The data are also oversampled in the manner of their analysis of the Mt. Blanc–gravitational wave data. One way to attempt to correct for the oversampling is by averaging the number of triple coincidences in a window ± 0.5 s about the maximum. Using the "peak" of Fig. 3 in Amaldi *et al.* (1989), we arrive at a "typical" 11 triple coincidences. If we use the correct expected number of 4.8 from the data set in question, then these have a Poisson probability of only 0.01 [see Eq. 2.7]. If we then correct for the freedom to shift the clock correction, for the introduction of this analysis method which was not used before, and for the narrowing of the data set to one hour (a correction of up to one order of magnitude), we see here as well that the correlations become completely insignificant. The claimed triple coincidences at other thresholds can be dealt with similarly.

Gravitational wave coincidences with KII and Mt. Blanc combined.

In the appendix, we criticize the fact that when the KII and Mt. Blanc data sets are combined and then analyzed in the manner in which the Mt. Blanc data set was, it is almost impossible to judge what is the independent contribution of KII to the resulting correlations. If the KII set does not exhibit the correlations independently, and we have just seen that it does not, then it cannot be expected to enhance the Mt. Blanc correlations.

We only wish to make one further remark about this: given the considerably greater sensitivity of Kamiokande to neutrinos than Mt. Blanc has, it is odd that RTM do not weight the KII neutrinos more heavily than the Mt. Blanc "neutrinos" in the combined data set. If RTM were to weight the KII particles more strongly, then their lack of correlations would, of course, depress the significance of the correlations RTM find in the combined data sets.

For all these reasons, we do not feel that the analyses of the combined data sets contribute significantly to the probability estimates we have been making.

Conclusion: Kamiokande as a test of the Mt Blanc correlations

We find no evidence that Kamiokande particles were correlated with the gravitational wave detectors in the way that Mt. Blanc "neutrinos" were. This is, as we have said, the acid test of the Mt. Blanc effect: if the correlation is not also present in the data of other neutrino detectors, then it is hard to believe that it is real. Despite the claims in the RTM papers about the existence of correlations in the KII data, the freedom they have to adjust parameters is enough to explain the weak correlations they find, on the null hypothesis that the KII data are random with respect to the gravitational wave data. *The RTM analysis of the KII data therefore provides the falsification of their hypothesis that the Mt. Blanc correlations are due to a real physical effect.*

2.5.3 Reassessment of gravitational wave–IMB correlations

The paper on the IMB correlations (Aglietta *et al.* 1991a) appears to have been written as an attempt to unify the derivation of the earlier correlations as well as to find new ones in the IMB data. In fact, this paper presents the Mt. Blanc, KII, and IMB correlations as a coherent whole, adopting a single (but new) dataset length of $1\frac{1}{2}$ h for each, and using the same delay, 1.2 s. Although these choices have been made *a posteriori*, it is still a welcome attempt to ensure comparability.

Using these new parameters, the significance assigned by RTM to the Mt. Blanc correlations is reduced by a factor of 30. The significance of the KII correlations is hardly changed for the following reason: in Amaldi *et al.* (1989), the delay examined is 1.1 s and the clock correction for the KII detector is estimated to be +7.7 s, where the effect is strongest. However, in the new analysis (Aglietta *et al.* 1991a), although they adopt a 1.2 s delay, they change the KII clock correction to +7.8 s, *thus cancelling the detrimental effect of changing to the new 1.2 s delay.* RTM say the effect of this small difference in the clock correction is "negligible", and they do not even point out their previous use of 7.7 s, referring to it simply as another clock correction they could have used. However, the effect is far from negligible: we have seen, as can be read from Fig. 1 of Amaldi *et al.* (1989), that this ad hoc 0.1 s adjustment makes a difference of a factor of about 10 in the strength of the correlation. We note also that RTM justify the new clock correction of 7.8 s because it matches the middle of the first 5 KII neutrino times with the middle of the first 3 IMB neutrino times. This vindicates our correction in Section 2.5.2 for the fact that RTM could have chosen methods of setting the clock correction other than the one they used.

Fortunately, the IMB clock is known to have been set correctly at the time of the experiment. However, in finding the correlation between the gravitational wave detectors and the IMB particle detector, RTM make three significant choices, all of which are *a posteriori* and for which no justification is offered.

Although in their earlier papers, they first emphasize that a 1.1 s delay is optimal (Aglietta *et al.* 1989), and they then refer to it exclusively as the optimal delay for the phenomenon, (Amaldi *et al.* 1989; Aglietta *et al.* 1991b) the delay is changed back to 1.2 s here without explanation. RTM do not state the "probability" of the correlation at 1.1 s, and nor is it given in any graph. It is hard to know how these changes affect their claimed probabilities; but we have seen, e.g., in Fig. 2.6, that changing the delay around a correlation can make a difference of a factor of about 10 in the "probability" of the correlation.

We must also consider RTM's decision to change the length of the dataset again, this time to $1\frac{1}{2}$ hours, still centred on 2h 45m. This, they say, is based on their previous analyses of two hours for Mt. Blanc and one hour for Kamiokande. Certainly, this is a good compromise between the periods analysed previously, and will not affect those earlier correlations too dramatically. But in fairness, RTM could quite easily have chosen a period of one hour or two hours instead, as they have used these before. In fact, they could have used any period of this order, as this is what they did when they chose to analyse one hour of KII data instead of two hours.

Finally, we must account for RTM's decision to apply an energy selection criterion to the particles detected during the experiment (see Appendix). They acknowledge that this choice affects the probability, but they make no attempt to justify it or to calculate the extent to which the probability is affected. Within the null hypothesis, of course, and particularly when these particles (high energy muons) could not possibly have been responsible for the effect observed in Mt. Blanc, which contained no muons (Aglietta *et al.* 1991a), this choice has no basis. It is impossible for us to quantify the effect of this choice, since RTM do not give any results for the whole IMB particle population.

In summary, we must correct RTM's *a posteriori* calculation that the probability of the IMB correlations is 10^{-3} for their adjustment of the delay; for the choice of particle sample; and for the choice of the period of analysis. We think that a factor of 100 for such choices would not be unreasonable, giving a probability for the correlations of as much as $\sim 10^{-1}$.

2.6 Conclusions

We have found that the Mt. Blanc "neutrino" data and the Rome and Maryland gravitational wave data streams show a weak correlation during the period of two hours containing the Mt. Blanc "neutrino burst". The correlation is of such a nature that it would have been found once in similar *a posteriori* analyses of between 10 and 1000 random data sets. In addition, we believe that the particle data from Kamiokande and IMB show no compelling evidence of the same correlation with the gravitational wave data, while RTM accept that there are no similar correlations between gravitational waves and the Baksan detector. This leads us to the conclusion that the correlations found by RTM are most likely a chance fluctuation in the data.

In reaching these conclusions we have had to try to compensate for a host of choices and other biases in the original RTM analyses. These include: *a posteriori* choices of the time-delay, of the threshold, and of the duration and starting time of the data set; statistical dependence of data sets caused by including the original "eyeballed" data set in the larger ones that were subjected to an analysis that was based on inspection of the original set; use of nonstandard and seriously flawed statistical tests with poorly-understood statistics, when standard tests could have been used but were not (or were not reported); and the failure to apply consistently the Mt. Blanc analysis methods to data from Kamiokande and IMB. In assessing the effects of some of these choices we have been guided by our own numerical simulations of the RTM methods applied to random data sets.

The result is that we believe that the correlations, while present, are very much more likely to arise in random sets than RTM estimated. Since the Kamiokande, IMB and Baksan data do not show the same correlations, any physical model for these effects would not only need new particles and interactions; it would also have to explain how the Mt. Blanc detector could have responded while the larger Kamiokande detector did not. We feel that the correlations present in the data are sufficiently weak that they do not provide serious evidence for such new physics.

Acknowledgements

We thank the following for useful conversations and correspondence: F. Dunstan, J. Hough, T. Niebauer, T. Piran, O. Saavedra, K.S. Thorne, and especially G.V. Pallottino and G. Pizzella. We also thank our referees for many useful comments and suggestions. CAD acknowledges support of a postgraduate studentship from the Science and Engineering Research Council.

2.7 Appendix: Review of the RTM analyses

2.7.1 The gravitational wave–Mt. Blanc coincidences

The first RTM analysis (Amaldi *et al.* 1987) dealt only with Rome and Mont Blanc data; the Maryland data were brought in later. On inspecting the raw gravitational wave and neutrino data near the time of the Mt. Blanc burst (see Fig. 2.2, which is reproduced from Fig. 2 of their paper (Aglietta *et al.* 1989)), RTM saw that a delay of 1.4 seconds between the two data streams would place the "neutrinos" in the Mt. Blanc event at times when the Rome signal is appreciably higher than the average background (Amaldi *et al.* 1987). Although this was unexpected, the time delay was at least consistent with a massive neutrino model with an acceptably small mass for the neutrinos. The probability that the unusually high signal preceding the first neutrino should have occurred within a time interval of 3 s was 0.03, not very significant in view of the large gravitational wave energy that would be required to explain it.

However, it motivated a much more involved analysis (Aglietta *et al.* 1989) in which the two sets of gravitational wave data are collated and examined for coincidences (see Section 2.7.2). The time under consideration in this analysis is the full 18 hour period, from 12h 22 February to 6h 23 February. During this period, the Mt. Blanc detector counted 775 events, 5 of which were in the "neutrino burst" at 2h 52m 37s. The remainder represent a fairly normal "background" counting rate.

In searching for coincidences between the two gravitational wave data streams and a neutrino stream, one might try the fairly standard approach of applying a *threshold* criterion to the gravitational wave streams, searching for coincidences among the three streams only when *both* gravitational wave detectors are above threshold. This would treat each data stream with equal weight; and moreover, if the data were random, each stream would be expected to produce "events" ("neutrinos" or gravitational wave threshold crossings) with a Poisson distribution of arrival intervals. The statistics of such a search would be easy to analyze.

RTM do not report having performed such an analysis on the Mt. Blanc data (although they did use this method later for analysing the Kamiokande data—see Appendix 2.7.3). Rather, their approach is first to combine the two gravitational wave data streams into a single one by either adding or multiplying them together, and then to use two different analysis methods—one of them a threshold criterion—to compare the combined stream with the Mt. Blanc neutrino data stream. We shall describe each of the two methods and their results separately.

Results of the net excitation method

RTM first examine the full 18 hours of data, calculating both $C_\times(-1.2\text{ s})$ and $C_+(-1.2\text{ s})$ for 2-hour stretches of data, moved along in $\frac{1}{2}$-hour steps. Each value is compared with 10^3 "background" values. They find the best correlation around 2h 45m UT on 23 February (Fig. 5, Aglietta *et al.* 1989), where $n = 1$ for $N_b = 10^3$. (For this value of N_b and 80 "neutrinos", there are about 5000 independent data values used for generating the 1000 background values, so the background data are probably independent in this case.) Then Eq. 2.4 would assign this a probability of 10^{-3}.

Importantly, this figure does not change by much if the 5ν burst is excluded from the "neutrino" data set: there are correlations at a 10^{-3} level with a time delay of 1.2 s even without the "neutrino" data that led to the suggestion of the delay. However, it is also significant that RTM use a delay of 1.2 s here rather than the 1.4 s used previously. It appears that they adjusted ϕ to get a better correlation.

To address the question of the best time delay directly, RTM next fix their attention on the 2-hour window about 2h 45m, and for the summation statistic $C_+(\phi)$ they vary the delay time ϕ, for each value comparing it with $N = 10^5$ to 10^6 "background" values. The tested values of ϕ are separated by steps of only 0.1 s, far below the gravitational wave sampling time of 1 s. The delays of 1.0, 1.1, and 1.2 s all have very low "probabilities" from Eq. 2.4, smaller than 10^{-5}. Importantly, the "probability" at 1.4 s is only about 10^{-4}, and within ± 0.5 s of the lowest point there are probabilities as high as 10^{-2}. (We have reproduced the RTM figure in our Fig. 2.6(b) above, where we compared it to the results of one of our simulations.)

Given that the gravitational wave signals are sampled at 1 s intervals, differences between time-delays as small as 0.1 s cannot have physical significance. The fact that the RTM "probability" changes by a factor larger than 10 when ϕ changes by 0.1 s gives us a measure of the confidence we can have in these probabilities. We can think of two possible causes of these large fluctuations.

First, the fluctuations may be simple small-number statistics: with typical ranking numbers of 10^5 or so, a given small change in C_* might lead to a relatively small change in n of, say, 10^3. But when we deal with points with ranking numbers near 1, a similar change in C_* would lead to a much larger *relative* change in n, say 2 or 3 or 10, leading to a much larger relative change in the inferred $p = n/N_b$. Thus, values of p substantially greater than $1/N_b$ are relatively stable against small changes in C_*, while the smallest values of p are fairly unstable. This is why we asserted in Section 2.3.1 that Eq. 2.4 could be unreliable if n is small.

The second possible cause of the fluctuations may be the non-independence of the background values. We distrust values of p near or below 10^{-4}, while the most significant correlations have values of p near 10^{-6}. It may well be that for these points the correlations among the background values ensure that there are tens of background values between any two real values of C_*, leading to spuriously large variations in p from one time delay to the next.

Given these problems, it would be more prudent to infer a probability from this method (if the method is to be used at all) by taking some sort of average over delays that span a 1 s interval. If the entire interval consistently gives ranking numbers less than, say, 100 for $N_b = 10^5$, then there may well be grounds for asserting that a correlation exists that has a probability of about 10^{-3}.

However, RTM do not do this: they consistently quote the lowest probability associated with any value of ϕ, even when very nearby time-delays have considerably higher values of p. This leads them to postulate probabilities as low as 10^{-6}. If a more prudent average were applied to the RTM data, the inferred probabilities would be in the range 10^{-3} to 10^{-4}.

When we discussed other problems with the RTM analysis in Section 2.5, we saw that further consideration of the free choices that RTM had in their analysis (such as the delay time ϕ) increases the probability of their finding such correlations in a random data set even further.

Threshold coincidence method results

By adopting a delay of 1.2 s and a threshold on the summation data stream $E_+(t)$ of 150 K, RTM find a substantial correlation. The number of gravitational wave events at this threshold is $N_{gw} = 172$. As before, there are $N_\nu = 96$ "neutrinos" and $N_t = 7200$ sampling intervals (2 hours), so the expected number of coincidences is $\bar{n} = 2.29$. The actual number observed was 13 (see Fig. 2.7(b)), much larger than the other values up to ± 50 s away, and with a chance probability (Eq. 2.7) of about 10^{-6}.

However, this figure does not take into account the fact that RTM (as they state) searched for coincidences at a variety of thresholds, and chose 150 K only because it gave the "best" correlation. This clearly affects any realistic assessment of the chance probability of this correlation, and provides one of the principal motivations for our Monte-Carlo study.

2.7.2 The Rome–Maryland gravitational wave coincidences

Although there were no improbable coincidences between the two gravitational wave detectors at the time of the Mt. Blanc event, the correlation with "neutrinos" over a long period of time makes it important to see if the two gravitational wave detectors were correlated with each other over this time. We give a brief summary of the RTM analysis (Amaldi *et al.* 1988).

Threshold coincidence method

RTM use data covering the 36-hour period from February 21, 18h 24m 33s to February 23, 6h 2m 3s UT 1987, which includes the Mont Blanc burst. During this period both detectors had good thermal distributions of noise. After rescaling the Maryland data by the mass ratio of the detectors (as described in Section 2.3.1), RTM set a threshold and count the number of times that both detectors are above the threshold simultaneously. The expected number can be calculated easily from the observed exponential distributions, and by calculating threshold-crossings with various delays one can test whether the data are behaving as expected.

Threshold coincidence method results

RTM divide the period under consideration into two smaller periods. Period 1, of about 7 hours (2.5×10^4 s) from February 22, 23h 5m 23s, to February 23, 6h 2m 3s, includes the Mt. Blanc burst. Period 2 is an *earlier* and longer period of 10^5 s, from February 21, 18h 24m 33s to February 22, 22h 7m 52s, which seems to have been analyzed as a control for the analysis of Period 1. RTM do not explain why they have chosen to place the division between the two periods at about 23h 5m 23s.

During Period 1 there was an excessive number of coincidences above a threshold of 100 K: 41 coincidences, which is 2.4 standard deviations from the mean. Although this threshold is arbitrary, RTM give an argument that the number is still large when accumulated over 26 values of the threshold from 25 K to 150 K, and that the correlation seen has a chance probability of 3.5%.

For Period 2, there is no significant correlation between the two antennas: 114, only 3 more than would be expected (the interval is 4 times as long as Period 1). RTM regard this as evidence that the detectors are behaving normally during Period 2, and hence, it is implied, during Period 1. (They do not, however, assess the probability that one would come so close to the expected value for this period.)

Given that the gravitational wave detectors show an unusually high rate of coincidence for the period that includes the Mt. Blanc correlations described earlier, it is important to ask whether the gravitational wave correlations can by themselves account for the gravitational wave–"neutrino" threshold coincidences during the same period. In other words, if "neutrinos" arrive randomly but the gravitational wave detectors are correlated (for whatever reason, even by chance), do we expect the number of "neutrino"-gravitational wave coincidences that are seen?

The answer is clearly no. The number of coincidences between the gravitational wave detectors above 80 K (giving a summation energy of 160 K) is less than 20 (2standard deviations) more than would be expected by chance in the 7-hour Period 1. During the 2 hours of the neutrino-gravitational wave coincidence analysis, this would probably give only 5 gravitational wave "events", and the chances of their being in coincidence with a random "neutrino" is very small. They could explain less than 0.1 of the observed 13 "neutrino"-gravitational wave coincidences at a summation threshold of 150 K. It seems, therefore, that the "neutrino"-gravitational wave threshold coincidences occur primarily when one gravitational wave detector is well below the excitation level of the other, and are probably not associated with the coincidences tested here.

It is not clear whether the net excitation correlation of gravitational waves and Mt. Blanc "neutrinos" is affected by the gravitational wave-gravitational wave correlation. It is possible that the small excess of gravitational wave coincidences at most thresholds and at zero delay will raise the value of $C(\phi)$ compared to the background, because the "signal" gravitational wave values are taken at the same time, whereas the background values are taken at different times. This may marginally increase $C(\phi)$ and hence decrease n.

2.7.3 The gravitational wave–KII–Mt. Blanc coincidences

Net excitation analysis of Kamiokande data

RTM first apply the "net excitation" analysis method to the Kamiokande-gravitational wave coincidence problem. For both the summation and product statistics, the best correlation occurs at a time-delay of $\phi = 6.6$ s, which they take to be composed of a clock-offset adjustment $\phi_c = 7.7$ s and an intrinsic time-delay of -1.1 s, consistent with the Mt. Blanc-gravitational wave time delay. They give a rough estimate of the chance probability of this correlation of about 10^{-3}.

Net excitation analysis of merged particle data sets

RTM next do something new. They merge the set of 105 Kamiokande particles in this 1 hour stretch with the set of 48 Mt. Blanc "neutrinos" during the same period, to give 153 particles in all. They apply the "net-excitation" method to this set, obtaining $C_{+}(\phi)$ (see Eq. 2.1), with results similar to those for the Mt. Blanc data alone. Using equation Eq. 2.2 to generate a background, they get $n = 3$ for $N_b = 2 \times 10^6$, so their probability estimate for this correlation would be about 1.5×10^{-6}. Of course, the Mont Blanc "neutrinos" will have contributed to this, so it cannot be an independent statistic.

RTM attempt to remove this dependence on the Mt. Blanc analysis by shifting the KII signals by large random times, thereby giving a control set, where the KII signals are certainly not expected to be correlated with the Mt. Blanc "neutrinos" or the gravitational wave signals. They recalculate $C_{+}(\phi)$ and n for $N_b = 2 \times 10^6$, and obtain $n = 23,942$, a higher value than that of the "correlated" data set, indicating a weaker correlation, which can only have come about from the Mt. Blanc "neutrinos". They take the ratio of these n values, viz. $3/23,942 = 1.25 \times 10^{-4}$, to be the experimental probability that the Kamiokande data's contribution to the correlation is purely chance. They then correct this for the fact that they have chosen the best value of ϕ_c from the 41 values they would have allowed themselves (steps of 0.1 s in the 4-second window for ϕ_c) by multiplying this probability by 41, arriving at a value of 5×10^{-3}, similar to the previous net-excitation probability.

It is not clear why taking the ratio of these two numbers should produce the probability of Kamiokande's "additional contribution", if any, to the Mt. Blanc–gravitational wave coincidences, and RTM make no attempt to prove this. In any case, the probability they arrive at is not very significant; but even here we must raise the same doubts as before that using steps of 0.1 s for time delays is unphysical when the data are sampled at 1 s intervals. RTM attempt to compensate for this correction by allowing for 41 independent choices for the time-delay, but it is not clear to us that this compensation is correct. We have argued above that they should instead average the probability values within a 1 s window, and this could give a much larger correction, since their method starts from the unphysically small value of C_+ at the best time-delay.

Triple coincidence analysis

RTM examine thresholds of 30 K, 40 K, 50 K, and 60 K for gravitational wave "events", on each detector. They do not explain why they choose these thresholds, which are much smaller than half of the threshold of 150 K that they adopted for the summation statistic in the Mt. Blanc analysis. Indeed, the triple coincidences become much less significant for the higher thresholds.

The expected number of triple coincidences, given a uniform distribution of arrival times, is

$$\bar{n} = \frac{N_\nu N_M N_R}{T^2} \tag{2.13}$$

similar to Eq. 2.6, with N_ν the number of particle events; N_M and N_R the number of Maryland and Rome gravitational wave "events", respectively; and T the length, in seconds, of the data set (here 3600). For the threshold of 40 K, this works out to be 4.8; for a threshold of 60 K it is 0.97. The actual analysis gives, for $\phi_c = 7.7$ s, 15 coincidences at 40 K and 5 at 60 K. The "raw" probabilities of these are, respectively, 1.5×10^{-4} and 3.2×10^{-3}, using equation Eq. 2.7 (RTM quote only the former, the smaller of the two). RTM would compensate for the freedom to choose a time-shift by multiplying each by 41, thereby producing numbers comparable with the earlier tests of the Kamiokande data. RTM go further than this. They calculate a Poisson probability for the number of observed coincidences, but for the expected number $\bar{n}$ they use the expected number taken from an *earlier* period of time! RTM seem to think that it is better to take the mean from the earlier data, since there is no suggestion that it contains real events, and they heavily emphasise the resulting smaller probability. But this is clearly not right: the Poisson probability formula [Eq. 2.7] only gives the probability of a given number of coincidences in a data set as against the expected number of events *in the data under discussion*, and using a mean from another data set will give an incorrect result. In fact, during the hour under consideration the number of Kamiokande particle events was some 25% higher than the hourly particle rate during the comparison earlier period. Therefore, RTM's contention that for the earlier comparison period, "the statistical properties of the data are very similar to those of the period of analysis" (Amaldi *et al.* 1989), is, in this sense at least, seriously incorrect. Furthermore, for the case of the 40K threshold, using the wrong expected number (3.7) leads to a probability estimate of 8.2×10^{-6}, as compared to 1.5×10^{-4} for the actual mean number (4.8).

RTM thus quote probabilities about a factor of 20 smaller than the correct ones. However, they clearly have some doubts about this calculation.

Merged triple coincidence analysis

Finally, choosing the offset $\phi_c = 7.7$ s, RTM again put the Mt. Blanc and Kamiokande particles together and search for triple coincidences between all the particles and the two gravitational wave detectors. When setting a threshold of 40 K, and comparing the numbers of triple coincidences in the 2h to 3h UT window of 23 February with the respective average numbers of triple coincidences taken from the period 12h to 24h of 22 February, RTM estimate the Poisson probability of the triple coincidences found to be 2.7×10^{-8}, again using a mean taken from an earlier data set. At other choices of threshold they estimate the triple coincidence probability to be between around 10^{-5} to 10^{-7}. This is again clearly dependent on the earlier Mt. Blanc–gravitational wave coincidences, though no attempt is made this time to correct for this. It also suffers from the fact that the number of Kamiokande particles is higher in the hour under analysis than in the comparison hours earlier. We do not see any way of using this analysis to estimate the independent contribution of the KII particles to the probability of the correlation.

Summary

Displaying some caution regarding these analyses, RTM prefer to adopt their earlier value of around a few times 10^{-4} (Amaldi *et al.* 1989) as their estimate of the significance of the independent support that the Kamiokande data give to the coincidences already found between the Mt. Blanc detector and the gravitational wave antennae. We have reassessed this claim in Section 2.5.2.

Note added after submission

After submission of this paper, it was pointed out to us by one of our referees that Aglietta *et al.* 1991b does give a clue as to what would happen if one performed a threshold coincidence analysis of the Kamiokande-gravitational wave data, for the original period used, viz 1h 45m to 3h 45m UT. From Table II of Aglietta *et al.* 1991b, there appear to be only 4 coincidences of Kamiokande particles with gravitational wave data above a summation threshold of 150K. This compares with an expected number of 5.0, obtained from: Eq. 2.6, from doubling $N_\nu = 105$ from the one hour period given in Table III of Ref. Aglietta *et al.* 1991b, and from $N_{gw} = 172$, given in equation 12 of Ref. Aglietta *et al.* 1989. If this were true, there would certainly be no correlation of the Mt. Blanc threshold coincidence type in the Kamiokande data, at above the chance level.

This is only what one can infer from the data presented in refs. (Amaldi *et al.* 1989; Aglietta *et al.* 1991b), which are slightly (but for these purposes, not seriously) incomplete. It would be helpful for RTM to publish the actual threshold coincidence analysis of the Kamiokande-gravitational wave data if they have not done so already.

2.7.4 The gravitational wave–IMB–Mt. Blanc coincidences

On receiving the IMB particle data, RTM analysed them together with the gravitational wave data and the Mt. Blanc and Kamiokande particle data (Aglietta *et al.* 1991a). For brevity, we shall concentrate here on the simple coincidences between the IMB particle signals and the gravitational wave data, since only these could provide an independent confirmation of the apparent Mt. Blanc–gravitational wave correlations.

Net excitation analysis of IMB data

RTM apply the same analysis technique, the net excitation method, to the IMB signals. This time they choose to analyse only the $1\frac{1}{2}$ hour period of data from 2h 0m to 3h 30m of February 23. RTM choose here to apply an energy selection criterion to the particle data. They define a quantity $E(IMB)$, the visible energy per event given by multiplying the number of photoelectrons detected during the event by 1MeV; 1MeV being, very roughly, the energy deposited in the detector by one photoelectron (Aglietta *et al.* 1991a). RTM then choose to analyse only those signals whose visible energies fall in the range

$$3 \text{ Gev} \leq E(IMB) \leq 6 \text{ Gev} \tag{2.14}$$

based on their inspection of Fig. 2.12 (which is Fig. 5 in Aglietta *et al.* 1991a). The peak in

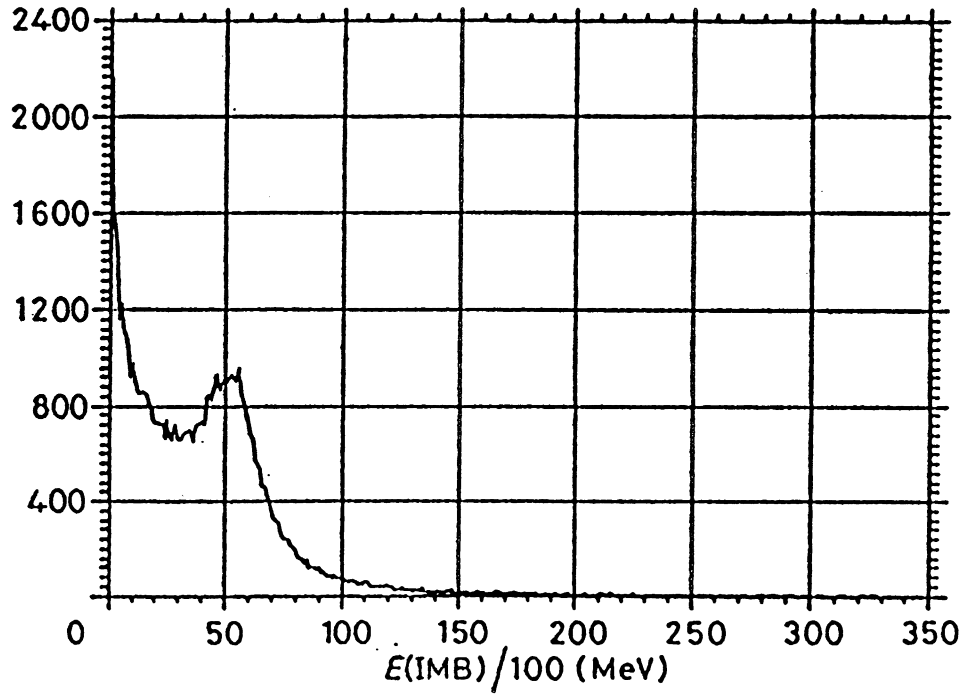


Figure 2.12: The energy distribution of the IMB particles during the period in question. RTM choose to consider only those between 3 GeV and 6 GeV. (Reproduced from Aglietta *et al.* 1991a with permission.)

Fig. 2.12 is due to single muons that cross the entire apparatus. This choice is very curious for the following reasons:

1. RTM have not previously made such a selection of particle detector data based on particle energies, although for the Kamiokande data RTM adopted the same threshold that is used by the Kamiokande group in their own analysis (Aglietta *et al.* 1991a);

2. the choice seems to suggest that the muons are responsible for the correlation, while RTM have not yet offered a consistent particle-based model for any of the correlations; the choice is therefore completely *ad hoc*;

3. the principle that new data from IMB should be subjected to the same analysis as that which found correlations in the Mt. Blanc data is somewhat compromised;

4. if the particles in Fig. 2.12 really do comprise two separate populations, and if one were particularly interested, *a priori*, in the "peak" population, the natural energy interval to examine would be that where *the "peak" population departs from the underlying curve*, i.e., between 3 GeV and about 8.5 GeV. RTM set their cutoffs where the peak departs from the underlying curve at the lower limit (3 GeV) and where the overall number drops back down to the level at the 3 GeV limit (6 GeV). RTM thus exclude those particles between 6 GeV and 8.5 GeV, even though they belong to the same population which RTM claim they are selecting! This criterion has no scientific basis, even were it correct to separate the two populations;

5. the Mt. Blanc data, where the effect was first seen, and where the effect is still claimed to be strongest, contain no muons (Aglietta *et al.* 1991a).

For the selected particles, and adopting a delay 1.2 s, RTM find a correlation whose probability of 9×10^{-4} is assessed in the usual way for the net-excitation method. The figure for the correlation at a delay of 1.1 s is not given. For an *advance* of 1.8 s, RTM find a correlation of claimed probability of 10^{-3}. We have reassessed this analysis in Section 2.5.3.

Chapter 3

Reassessment of the Reported Mont Blanc–Baksan Neutrino Coincidences[1]

An analysis was performed on the Baksan and Mont Blanc neutrino signals by Chudakov (Chudakov 1989). Chudakov claims some significant correlations; although he personally believes that, if they are real, they were not caused by neutrinos and were probably not associated with the supernova (Chudakov 1989).

The reader must bear in mind that RTM have analysed the Baksan data for correlations of the type found before, and *RTM state explicitly that there are no correlations of the Mt. Blanc type in the Baksan data* (Aglietta *et al.* 1991b). The analysis which RTM actually do on the Baksan data is of a completely different type, and at this stage that analyis is worth little as any kind of support for the Mt. Blanc effect. But because Chudakov's work was done more in the spirit of the original analysis, using the correlations in the same way, even though it only involves "neutrinos" and not gravitational waves, it is worthwhile to look at the data to see if there is any hope of an effect.

In this chapter, I shall present Chudakov's analysis and his inferred probability of the correlations found; and then I shall present my own reassessment and my own estimate of

[1]Note: This chapter functions as an addendum to Chapter 2, as it analyses more claimed coincidences between particle detectors at the time of the "Mt. Blanc burst". It was not included in the original paper for reasons of increasing length of said paper, and because it makes little explicit reference to gravitational waves: the coincidences claimed are mostly between particles only.

the a *priori* probability of the correlations. I shall conclude that the Baksan data analysed in this way provide no convincing new evidence of a new physical effect observed on 23rd February 1987.

3.1 Chudakov's Assessment of Mont Blanc–Baksan Coincidences

Again, there is a major problem in the setting of the Baksan clock, which is open to uncertainties of between -54 s and $+2$ s. Chudakov postulates, following RTM, that the first neutrino of the Baksan burst at 7h 35m should be coincident with the first neutrino of the IMB burst, making use of the fact that the IMB clock was well set. This gives a correction of -30.4 s to the Baksan clock, except for an uncertainty which is unknown, but which he guesses as being anything from ± 1 s (which Chudakov calls "most optimistic") to ± 5 s ("conservative") to anything between -54 s and $+2$ s ("most conservative" i.e. this method of setting the clock doesn't work). Since we have seen (Chapter 2) that the length of the IMB burst at 7h 35m 41s was 5.6s long, while the corresponding Baksan burst was 5.7s, it seems fair to us to allow a "reasonable" error to be up to $\sim \pm 5$ s.

Chudakov uses this choice of clock adjustment, and finds that there is a number of coincidences between Baksan and Mt. Blanc, in a window of ± 1s, in a one-hour period centred on 2h 15m whose poisson probability he estimates at 2×10^{-4}.

Chudakov also finds an excess of triple coincidences between Baksan, Mt. Blanc and the gravitational wave signals, where, the reader must assume because it is not stated explicitly, the threshold on the gravitational wave sum was the same as before; i.e. 150 K. These were found in the same time period as above, 1h 45m to 2h 45m, and are assigned probability ~ 0.03.

3.2 Reassessment of Mont Blanc–Baksan Coincidences

Again, as with RTM, Chudakov has changed several parameters from the original RTM analysis, some of which didn't need to be changed. It is hard to say, in some cases, whether this was done to improve the correlation and, if so, by how much. We need to correct back, if we can, to RTM's original choices to see the a *priori* probability of the correlations between

Mt. Blanc and Baksan, and whether they confirm or deny the Mt. Blanc correlations.

3.2.1 Choice of dataset

The dataset was definitely changed to improve the correlation, and Chudakov says narrowing the dataset to one hour lowers the probability by a factor of four. This is unacceptable in view of what we have seen already — one is not allowed simply to vary such parameters as the length of the data set as one feels. The a *priori* choice should be the same as the experiment on which Chudakov's analysis is based, i.e. a conference talk based on Aglietta *et al.* (1989) cited in the *Correlations Reassessment* paper (Chapter 2) , where the length of the data set is 2 hours. Consequently, all probabilities found by Chudakov in this one hour of data must be adjusted up by a factor of four.

But Chudakov also *moves* the dataset. Although RTM centre their dataset on 2h 45m in the Mt. Blanc–gravitational wave correlations, and then centre it on 2h 30m for the KII coincidences, Chudakov centres his dataset on 2h 15m. In fact Chudakov's dataset doesn't even include the Mt. Blanc "neutrino burst" at all![2] He does not say explicitly what is the effect of moving the dataset. But one can see from his Fig. 1 of Chudakov (1989) that if the dataset were centred on 2h 45m, which is RTM's original choice, the probability of the correlation would shoot up to 10^{-2}; a worsening of a factor of 50.

In moving the dataset, Chudakov's only motivation is this figure; i.e. he chooses the best dataset (Chudakov 1989). This is after he has cut the length of the data set down to 1 hour. Therefore I must conclude that the correction to the probability for changing the dataset is something like 4×50, or a factor of 200.

3.2.2 Clock correction

One of Chudakov's choices was forced, i.e. the clock correction; but this will not necessarily prevent its being a free parameter. Chudakov claims that he uses the same method as RTM used for the KII analysis, in claiming that the first neutrino of the Baksan burst at 7h 35m should be simultaneous with the first of the IMB burst, the IMB clock being accurately set.

[2]Obviously, if Chudakov wanted to stay close to RTM's model, scant as it is, he should at least have included the Mt. Blanc burst, which RTM claim is what led them to their choice of time delay. Thus Chudakov, in not following their model, has given himself great freedom. He actually chooses the best 1-hour period in his graph of 1-hour periods, which goes from about 0h to 6h UT. (Chudakov 1989)

However, it is hard to say from the paper whether his method for correcting the clock was chosen primarily because it was the same method as used by RTM (an approach which I believe would be the best under the circumstances), or whether the method was adopted after he had searched and found that the best correlation occurred close to the offset which would be given by that method. If one is tempted to give Chudakov the benefit of the doubt, remember that he has *not* followed RTM in the choice of the length or location of the dataset. In fact, the only way he has followed RTM's method is that he has adjusted the dataset to give the best correlation! We must also be sceptical, therefore, about his choice of clock offset, in that he would have had no problem in justifying clock offsets which were up to 5 s different to the one he uses.

In fact, even although the best data set has been chosen, the effect is well reduced outside a ± 1 s window about the optimum delay time, and is completely gone outside ± 2s (seen from Fig. 2 of Chudakov 1989). Chudakov would be very optimistic indeed to claim that the clock can confidently be corrected to this accuracy simply by imposing the criterion that the first signals of Baksan and IMB should be simultaneous (which is a claim he does not make), so one must correct Chudakov's probability estimate for the available choices of delay. (Chudakov himself makes an attempt at this.)

Also, while the resolutions of the detectors are ~ 0.1 s and the delay is shifted by this amount, the acceptance window is ± 1 s.[3] Therefore, when deciding how to correct for the choice of clock setting, it is not clear whether a choice window of ± 5 s is 10 or 100 choices.

Although there are only 10 completely independent choices of offset, there are 100 which are partly but not wholly dependent on each other (even within the best 2 s window, the strength of the effect does vary), and Chudakov chooses the best one of these 100. It isn't obvious what the correction factor for this should be, but it will be somewhere between 10 and 100. Since the probability which Chudakov infers only varies by a relatively small amount (a factor of three or so) within ± 1 s of the best delay, I would guess that the correction would probably be closer to a factor of 10 or 20.

[3] Although the data are being added over bins of 2 s, the particle detectors have resolution down to 0.1 s, so adjusting the delay to within 0.1 s is not oversampling in the sense that the gravitational wave data were oversampled in Aglietta *et al.* (1989) (see Chapter 2).

3.2.3 Delay window

Chudakov also takes the acceptance window for coincidences to be ± 1 s, different from the earlier RTM window of ± 0.5 s, and he seems to gloss over this difference. I do not understand why Chudakov has chosen a ± 1s window of acceptance, as opposed to RTM who always use ± 0.5s.

It is worth pointing out that Chudakov is not widening the window because Baksan is farther distant than the other detectors. The distance between the detectors is much smaller than the distance to any astrophysical source, so the time of arrival will not change significantly. In any case, it is no farther from Mt. Blanc than is, for example, Maryland.

However, I do have the feeling, in reading the paper, that this choice was made a priori; and Chudakov says that at least two other window widths are better (so he has searched windows other than the one he uses). Certainly, I would prefer that Chudakov had stated explicitly one way or the other whether this made any difference. But even though this would seem not to have been adjusted to improve the correlation, we are still primarily interested in the strength of the effect where all parameters are as similar as possible to the original analysis, i.e. for a window of ± 0.5 s. However, in the absence of a statement or a graph indicating whether this choice of window makes the correlation stronger or weaker, I cannot correct the probabilities obtained.

3.3 Corrected probability

In summary, therefore, I must correct Chudakov's assessment of the probability by a factor of 200, for the optimal choice of dataset, and for a factor of about 10 or 20 for the choice of clock correction. Chudakov assessed the probability of the correlations, given all his choices, as being 2×10^{-4}. My reassessment is that if one tries to make the dataset used as similar as possible to that originally used by RTM, and if one corrects for the freedom of choice of the time offset, the a *priori* probability of the correlations rises to around 0.4. Thus, there is nothing in the Baksan data, as presented by Chudakov, which supports the effect seen in the Mt. Blanc and gravitational wave data.

Finally, Chudakov finds some triple coincidences between Baksan, Mt. Blanc and the gravitational wave antennae, having probability ~ 0.03. One could argue that the same correction to the probability, a factor of 200, will apply here as before, since the dataset is the same and the analysis is similar (both analysis methods claim some sort of correlation between Baksan and Mt. Blanc, although the latter analysis method also involves some of the gravitational wave data). However, we must recall what we conluded in the *Correlation Reassessment* paper (see Chapter 2), that different analysis methods measure different, albeit related, properties of the data. Thus the correction may not be so much. Without more information from Chudakov, it is difficult to infer exactly what the correction should be, since he does not state how many triple coincidences one finds in the original 1h 45m UT–3h 45m UT dataset. Hence, the a *priori* probability of these triple coincidences could be anywhere between 0.03 and chance level.

3.4 Conclusions for Mont Blanc–Baksan Correlations

Thus I find that there is little or no evidence for RTM-type correlations between the Baksan and Mt. Blanc particle detectors. One of the methods (a method based on counting coincidences in a window, corrected for clock uncertainty) gives a result at chance level when analysed in as similar a way as possible to the original Mt. Blanc–gravitational wave correlations. Also, the correlation of triple coincidences one finds between the two particle detectors and the sum of the gravitational wave signals is very weak, between 0.03 and chance level. Both claimed correlations are based on a dataset which does not even contain the Mt. Blanc "neutrino burst" at 2h 52m on 23rd February 1987. This must also be added to the fact that RTM say that using their own analysis method, there is no evidence for Mt. Blanc type correlations in the gravitational wave–Baksan data.

It must be said, in fairness, that the energy threshold for Baksan is 10 MeV, nearly twice that of Mt. Blanc (Chudakov 1989). Therefore, one could devise a model where the particles which would have been responsible for the Mt. Blanc correlations did not excite the Baksan detector. However, this would necessitate further complications in an already contrived model required to consistently explain all the observations, on the premise that the correlations found are due to a real physical effect. Meanwhile the Baksan analysis presented here provides no convincing new evidence for such an effect being real. This makes it even more likely that the correlations seen between gravitational wave detectors and the Mt. Blanc neutrino detector on 23rd February 1987 arose by chance.

Part II

Chapter 4

Lessons Learned for Coincidence Analysis and Pioneering Research at Low S/N

In this chapter, I will try to outline the lessons learned from the RTM business, so that it will (a) serve as a warning for the near future of this nascent science, and (b) guide me in my own coincidence analysis of the 100 Hour Coincidence Experiment between two prototype interferometers (see Part III). I hope that this chapter will not be too negative or polemical, but rather that it will help to point towards a rigorous and secure future for the science.

Sections 4.1, 4.2, and 4.3 of this chapter will contain some discussion on various aspects of RTM's analysis. Section 4.4 contains a summary of the main points, and how they relate to my coincidence analysis of the data from the 100 hour experiment, to be addressed fully in Part III.

4.1 Post mortem on the RTM analyses

In my opinion, RTM's analysis firstly shows how not to analyse gravitational wave data and, indeed any data at low signal to noise, where there is ambiguity as to whether a signal is present or not. In RTM's case, since the data were gravitational wave data and particle data, one from a nascent science and one from a very young science, there was always the possibility of some unpredicted source(s). History abounds with examples where other young branches of astrophysics have found new unpredicted objects, and I am sure that gravitational wave astronomy will be no exception. Thus, with gravitational waves, there is a freedom to find and identify new sources. However, this freedom is a curse. And especially where there is so much riding on the first confirmed detection of a gravitational wave (money, prizes, fame) and where the signal-to-noise ratio of *any* sources, modelled or otherwise, is likely to be so low in the early stages, there is a great temptation simply to get carried away.

"Chase the signal"

RTM's approach was to "chase the signal", i.e. to adjust parameters from one dataset to the next, in the belief that so doing was emphasising a real signal which was already there. I believe, however, that all they were doing was optimising random correlations in each dataset; correlations which, when one corrects to their original parameter choices, wash out dramatically.

The "chase the signal" approach should, when possible, be reserved for when one is already fairly sure that a physical effect is responsible. The best way to be sure of this is to stick to the null hypothesis until one is forced to reject it; and to vary *as little as possible* from one experiment to the next, including the time of the effect, the method of analysis and so on.[1] If a signal is confirmed with high confidence, *then* is the time to vary the parameters of the experiment (time, threshold, etc.) so that more can be learned about the physics of the source.

It is not fair, however, to accuse RTM of adjusting everything they could to achieve the best correlation. The fact that there was such uncertainty in the clock timing, for example, was not of their construction; and there was little they could do *but* vary the time delay, given that they wanted to try to find a correlation. Unfortunately, this single uncertain parameter was so badly set in two of the particle detectors that, even if there were real

[1] By the same token, one should not vary parameters while attempting to contradict another experiment. In the case of the gyroscope experiment of Hayasaka and Takeuchi (1989) referred to later (see Section 4.3), this important rule was not followed: the "null result" repetition of the experiment by another group (Faller *et al.* 1990) was compromised by the fact that their second gyroscope had a mass, diameter, and material composition which were all different to the first. Hence, the experiments were not necessarily contradictory.

If one is tempted to think that material composition shouldn't make any difference, remember that unless one has a theory or model for the observations of Hayasaka and Takeuchi (1989), one must allow for some apparently bizarre new physics to be involved.

correlations present in the data due to real signals of some sort, RTM would probably never have found them with much confidence. With uncertainties so bad in the clock setting, RTM couldn't really win. Of course, RTM should at least have kept the other parameters fixed from one dataset to the next.

Obscurity of RTM's analysis

The preceding complaint can, at least in some cases, be cured; as B.F. Schutz and I attempted in our *Correlations Reassessment* paper (Chapter 2). However, my biggest objection to RTM's method is that many of the analyses which they do are obscure; in some cases, apparently deliberately so. Their analyses are obscure in two ways: (a) The analysis methods and ways of calculating the probability of certain correlations were frequently non-standard and difficult to assess: it is not obvious whether all of the small numbers they obtain are really the probabilities of the correlations occurring by chance; and, if they are genuine probabilities, whether they are are robust. (b) The analyses are spread across many journals and publications; where, on many occasions, the results and the analysis methods are changed *even sometimes for the same datasets.* For these reasons, it is difficult to find errors which may or may not be there.

The first duty of a scientist is to be clear about his method and findings, so that they are *testable.* RTM's failure in this is demonstrated by the amount of time and work which it took to find all the analyses, or small changes thereof, which I would describe as questionable: there shouldn't be so many and they shouldn't be so hard to pin down.[2] When dealing with such potentially important new results, RTM have a duty to be clear. Any obscureness in their analysis must be seen as a failure of their method.

Furthermore, I have indicated in a couple of places, in the *Correlations Reassessment* paper, where I find RTM's presentation of their work very misleading. For example, in the paper which analyses the IMB particles, and attempts to present the correlations observed in all the detectors as a coherent whole (Aglietta *et al.* 1991a), they exclusively select for analysis only a subsection of the IMB particles observed: those which they believe are the "high energy muons crossing the entire apparatus" (Aglietta *et al.* 1991a). However, in the analysis on the Mt. Blanc particles in the same paper, and in other analysis on the Mt.

[2]In fact, RTM have published other papers with other analyses which I have not even attempted to examine in this thesis.

Blanc particles (Aglietta *et al.* 1991b), RTM state with reference to the Mt. Blanc particles, "the signals due to muons had been eliminated before these data analyses." Thus, RTM use two mutually exclusive particle sets from the two detectors and analyse them side by side to obtain correlations which are meant to corroborate each other. RTM claim "In conclusion the new data (Kamiokande and IMB) exhibit a correlation with the GW detector data that shows up with the same characteristics and at the same time of that already found with Mont Blanc." But how can the characteristics of the correlation in IMB be the same as that in Mt. Blanc, when RTM have used mutually exclusive particle sets from each detector? Indeed, how can the same physical effect be responsible for the correlations in Mt. Blanc as in IMB, when RTM admit that they see no correlation in the IMB data unless they use only a type of particle which was excluded from Mt. Blanc (Aglietta *et al.* 1991a)? This striking anomaly in RTM's analysis method suggests very strongly that the Mt. Blanc correlations and the IMB correlations cannot offer any sensible corroboration for each other. That RTM should make such a strange particle selection is surprising. Most surprising of all, however, RTM make *no explicit comment* on this anomaly, in the paper in question.

Also, in the same paper, RTM change the time delay slightly from a previous paper which dealt with KII correlations, in order to improve the correlation in the other three detectors; but they also change the clock correction of KII by the same amount in the opposite direction, thus retaining the optimum time offset for KII. This is done quite deliberately to improve the strength of the effect, and RTM must know that it improves the strength of the effect, since it can be seen from their own earlier paper that it makes a difference of *about a factor of 10* in the strength of the correlation. However, RTM again make *no explicit reference* to their previous analysis, and worse they state "the effect of this small deviation on the analysis is, of course, negligible." This is not true, and at least someone involved in the writing of the paper must surely have known that it was not true.

One has to wonder what else RTM have not told us.

4.2 Models: working with and without them

4.2.1 Analysis of noise

The standard model adopted for most statistical analysis of very noisy data is that (a) current, well-agreed theories are correct, and (b) the data are pure noise. This is known as the *null hypothesis*. When these assumptions lead one to highly improbable conclusions, it is appropriate to question the null hypothesis.

Standard methods of analysis

Clearly, when one has decided on the null hypothesis, one has to be sure that the noise is analysed correctly, and that any probabilities assigned to the data are correct. One way to do this is to keep the analysis methods as standard as possible, i.e. the statistics and distributions of the tests should be well-known and robust. To adopt *ad hoc* and untried analysis methods may give probabilities which appear reasonable, but are actually of doubtful robustness (e.g. RTM's ranking order of the function $C(\phi)$), or which, in some cases, are just plain wrong (e.g. RTM's calculation of the Poisson probability of a number of coincidences based on a mean from a different period of time).

Furthermore, to adopt obscure analysis methods, as RTM did, invites suspicion, whether deserved or not, that the standard analysis methods were tried and the desired result was not found. Fair analysis should not only be done, it should also be seen to be done.

When the standard methods are unsatisfactory

Where the standard analysis methods are, for some reason, unsatisfactory, one could try to create a new analysis method. For instance, if RTM really believe that their analysis method, that of summing gravitational wave bar detector output at a certain time and comparing it to a background of sums with time offsets between the detectors (Pizzella 1988; Aglietta *et al.* 1989), really may help to expose signals, despite the detectors being misaligned and having different resonant frequencies, then it is reasonable to try this new method. Of course, one must find a way of testing the reliability of such a new method. One could try, as B.F. Schutz and I did in the *Correlations Reassessment* paper, Monte Carlo-simulated data, analysed in the same way as the real data. In fact, when the Rome and Stanford bar groups recently performed a coincidence analysis on the output of the two bar detectors (Astone 1992), they each provided the other with a real dataset and a "placebo" dataset, which came from another time (they called this a "double–blind" analysis; although it is, in fact, a simple Monte Carlo test, with one control dataset). This is a fairly satisfactory way to proceed, unless the "placebo" is identified as such, which is what happened in that case (Astone 1992).

4.2.2 Working models and the null hypothesis

It is normal practice in physics to have some sort of working model, which explains all the current theories and observations in a coherent or semi-coherent whole. Of course, no two scientists will agree on this model in every detail. But in order to perform data analysis, one has to adopt various general a *priori* assumptions; for example, that general relativity is correct, that burst sources are distributed uniformly throughout the galaxy, and so on. Every now and then, however, a result comes along which cannot reasonably be explained in terms of the old model, and one is forced to question it or reject it. The old model is questioned on the grounds that the result is so unlikely, were the old model true and were the data pure noise, that the data were more likely to have had the said outcome if the model were not true.

However, this tacitly supposes something, which RTM never address, to my knowledge: that *there exists another model which, if it were true, would result in the data outcome being more likely.* In fact, RTM never give a consistent model for their observations; and they never address the possibility that, even if the correlations were as unlikely as they claim, there may not be a model other than the null hypothesis which explains the observations better. One could argue that if no new model can consistently be constructed, wherein the observations were more likely to arise, one must retain the null hypothesis no matter how unlikely were the correlations found. (This would not preclude, of course, publication of said correlations with no model; which is what RTM did.) I believe that RTM attempted and failed to construct such a model, and this is why they never suggest one in the literature. Until they do this, in my opinion, RTM cannot legitimately claim, as they do (Aglietta *et al.* 1991b), that "real effects have been observed in independent, different and at intercontinental distance detectors during the supernova SN1987A."

In fact, even if all the correlations which RTM find were as strong as they claim, a *priori*, it is doubtful to me whether any self-consistent model *could* explain the observations better than that they had occurred by chance. I am led to this conclusion (1) by RTM's particle selection, where they select mutually exclusive particle samples from different detectors, (2) by the fact that the gravitational wave detectors were not aligned nor were of the same resonant frequency, (3) by the fact that, due to the small detection probability per neutrino in the Mt. Blanc detector if neutrinos were responsible, and due to the high background count rate, most of the family of particles which RTM claim were responsible for the effect almost certainly came from the background count, and (4) on energy grounds: if the particles or interactions responsible for the effect were emanating from the supernova in the Large Magellanic Cloud, the large distance to the source, coupled with the small target area of the detectors, coupled with the long period over which the phenomenon was observed, would point towards there being much more energy released than could be contained in one star with the mass of a few tens of solar rest masses (remember, RTM claim the optimum energy threshold for the gravitational wave detectors observing the phenomenon was 150 K in summation: Aglietta *et al.* 1989).

Even if such a model could be constructed, for Bayesian purposes one must also consider the "likelihood of the new model being true". RTM's required new physics fares badly here: partly on aesthetic considerations, i.e. that the new model would have to be very bizarre (if it could reasonably be constructed at all), and partly on the grounds of my belief that, if some new physics were involved, one may have to throw away much old physics which has been well-tested (e.g. general relativity).

The general question remains, then: What does one do if the the observations couple with sufficiently low probabilities to indicate that one should adopt a new model, or that one has found some new kind of source?

4.2.3 When to change the model

The answer to this question is, not surprisingly, that "it depends". To take a hypothetical example, suppose that the LIGO interferometers are working at their best sensitivity. Suppose also that some events in the data are very suspicious-looking, such as regular pulses of short duration, but such a source has not previously been modelled. Suppose further that there is no obvious electromagnetic counterpart. How would one decide whether this is a real source?

The answer lies partly in one's ability or otherwise to ascertain whether the source is astrophysical, and partly on exactly how low is the probability of the data arising from noise. In this case, if the pulses went on for a long time, and if there were a real astrophysical source for the observations, one should expect to see diurnal and annual doppler changes in the frequency and time of arrival of the source, assuming the source were far enough outside the solar system. (This is essentially what happened with the first discovery of a pulsar; although, in that case, one was dealing with a very much higher signal-to-noise.) If one also had three or more detectors, one should even see the source move across the sky. This is easy. What if matters were not so easy?

If the phenomenon recurs only infrequently, but with low probability

If, however, our pulsating source did not last long enough for these considerations, or did not recur often enough, there are other avenues towards the ascertaining of its origin. For example, one could take a Bayesian approach, involving two stages:

1. assign a prior probability estimate to the likelihood that such a source exists (and does not have an observed electromagnetic counterpart);
2. calculate the a *posteriori* probability with which such a waveform would arise in noise.

Stage 1 is clearly unsatisfactory, because of the arbitrariness of assigning a probability to the Universe's containing a certain kind of object never seen before. But one always has to attempt to correct for one's ability to set a free parameter; in this case, the signal profile. One way to solve this problem is to argue along these lines:

In our example above, that of a pulsating source, the waveform is easily noticed by the observer's eye. Now, of all the possible random associations of data into various signal profiles, there is only a very small subset thereof which would have, to the human eye, a discernible pattern.[3] Intuitively, then, such a waveform seems much more likely to come from a real physical effect than an apparently random jumble of noise would. Further, it is not unreasonable that such a source could exist, particularly if theorists could construct a source object emitting gravitational waves in this fashion. In this case, since the gravitational wave data had a low probability of occurring by chance (since they had high enough signal-to-noise ratio to be noticed) it would not be unreasonable to *postulate* this as a new source. Given this as a "template" for future observations, any further detections would be a *priori* in nature.

Thus, in the light of physical and statistical evidence, our Bayesian calculation has boiled down to a simple a *priori* one; and a threshold with false alarm rate of, say, 1 event per year

[3] In general, the set of regular repetitions, straight lines, dramatic jumps etc. in data form a very small subset of all the possible combinations of random noise samples. This is similar to the argument which suggests that there must an arrow of time, because a mirror is more likely to break into an apparently random jumble of glass than a jumble of glass is to fall into the shape of a mirror.

could be set on these sources in the normal way.[4] More explicitly: the low chance probability with which those waveforms already observed should occur, coupled with a physical model to explain the emission of such waveforms of gravitational radiation, indicate that our prior Bayesian guess as to the likelihood of this source existing should be assigned a high value, of order 1.0.

However, since these data have been used to construct the model that we wish to test, they should not then be used to calculate the a *posteriori* probability of the signals arising in noise: once we have decided on the model, we must wait for fresh signals of the same type.

It is unfortunate that this method involves sacrificing the signals which led to postulating the new source. However, if such a new source type is to be a viable object of interest, it must have a reasonable event rate of, say, one every few years. Therefore, we will only have to wait about this long to confirm the source. If the claimed event rate were much less than this value, (a) one would need to postulate further that we have been very lucky in observing such a rare object in the short detector operation time; and (b) one would need to wait a long time for the next such event, in order to confirm it as a genuine new phenomenon.

An interesting analogy to this is the claimed solar neutrino–solar activity correlations, due to Gavryuseva, Gavryusev & Rosljakov 1990, referred to later in this chapter. The proponents claim that the probability of the correlations they find is less than 4%. Furthermore, if one adjusts the energy threshold (!), the correlations become more unlikely still (this was stated at the Texas–ESO/CERN Symposium in Brighton, 1990, but I have no written reference). Does this point towards an interesting new physical process?

Perhaps. Alas, we cannot admit the threshold-adjusted probability as evidence, for reasons similar to those in the case of RTM. However, as I have said elsewhere in this chapter, this adjusting of the threshold may be interesting to help us understand any new physics if and when new physics becomes necessary to explain the observations. The important result is the a *priori* one, of a few percent. Again this, on its own, is not very convincing: clearly, hundreds of important physical experiments are carried out every year, and some of them will inevitably give results which by chance have an a *priori* probability of a few percent. However, in a fashion which is similar to what I have said about claimed new sources of

[4] In general, the false alarm rate of a certain type of event should be less than the expected event rate. Otherwise, we will have as many or more false alarms as we have real signals. The expected event rate for a new phenomenon must be calculated on theoretical grounds.

gravitational waves, all we have to do is wait. If real physics is at work here, the correlations should become stronger with time. This is particularly true if the threshold of interest is retained: in another 10 or 20 years the probability should drop again; and this time the arguments against adjusting the threshold would fall, because the threshold will have been set *a priori*.

What I am saying, then, is that if one finds unusual signals or correlations, commensurate with a new type of source object, one should not quote the probabilities calculated in the pinning down of the source as if they were *a priori*: these probabilities will depend on how the parameters of the experiment have been adjusted to make the source stand out. One should, in fact, use those observations to set the parameters for any future experiments. The next time the effect is seen, it will be an independent and *a priori* confirmation of the claimed source. This is something that people will believe.

4.3 Low S/N analysis in physical sciences

The dispute of important new results

It is an observed fact of human nature that, when a scientific experiment of almost any type gives a result which is unusual or unexpected or difficult to fit into the current world view, those who claim the result are usually much more enthusiastic about the effect being physically real than are their peers in the scientific community.[5] The correlations claimed by RTM fit well within this pattern. One could also cite recent examples of quasar alignments (Arp 1987), redshift periodicity of galaxies (Tifft 1976; Guthrie & Napier 1991), temporal correlations between solar neutrino capture rates and solar activity (Gavryuseva, Gavryusev & Rosljakov 1990), anomalous weight reduction of a gyroscope (Hayasaka & Takeuchi 1989), etc.

This may be partly due to a (jealous) hope, on the part of the researcher, of discovering something radically new or exciting; and equally a resentment among his peers that he shall pre-empt them in doing so. The fault may also lie, in part, with the inflexible scientific method, and with inflexible scientists. It may also be partly due to general credulity on the part of some researchers; and there are even some who seem to make a career out of such

[5] Conversely, there must be many researchers who obtain results which are contrary to their own opinion or to the current model, who do not publish. This is another problem.

apparently extraordinary claims. I am certainly not going to condemn a generally open-minded approach: I believe that science progresses as much or more through the infrequent huge leaps forward in understanding, which are often opposed even if they later turn out to be correct, as it does through the normal steady plod to which most of us are condemned. Science needs ideas, crazy or otherwise, and it needs heroes. *But there is no excuse for not being thorough with one's method and conclusions.*

I believe the above examples are demonstrations of one of the most basic tenets of the scientific method, that extraordinary claims require extraordinary evidence. This is as much true for gravitational waves as for any other branch of physical science. In the case of RTM, the claims were extraordinary but, I believe, the evidence was not. This was made worse by the fact that the experiment was not repeatable, which contrasts, for example, with the gyroscope experiment.

The importance of the repeatability of an experiment

This non-repeatability was why it took so long for meteorites to be accepted as being of astrophysical origin. Largely until E.F.F. Chladni's seminal work in the late 18th and early 19th centuries, the scientific community would rather question the sanity and integrity of the meteorite finders than question the model: that there were no stones in the sky, therefore stones did not fall from the sky. In the case of gravitational waves, even though we already have a theoretical model for their existence, not to mention indirect evidence from the pulsar PSR 1913+16, perhaps the community will only be fully convinced of the first detection if this conclusion is unavoidable. This will be so if (a) the experiment is repeatable (such as continuous wave sources), or (b) the detection is very high signal-to-noise *and* coincides with simultaneous electromagnetic observation of the same source, electromagnetic observations being generally much higher in signal to noise than either gravitational wave or particle observations (as well as being older and more trusted than the other two relatively new sciences). This presupposes that the observation falls within the current model, which makes the evidence required less extraordinary.

4.4 Lessons learned for gravitational wave coincidence analysis

One of the most important general points to be learned from RTM's analysis is that coincidence analysis is a powerful tool; and the results can be powerfully misleading, if this tool is abused. Some more particular points, learned from this chapter, will specifically influence the analysis of the 100 hour data in the following ways:

1. In part III, I shall not, where possible to avoid, make a *posteriori* adjustments of parameters. Where data selections are necessary (e.g. see Chapter 7), these selections shall be performed "blind" to consequences for experiment sensitivity or gravitational wave detection. That is, I shall not "chase the signal".

2. My analysis shall be as clear as possible. All tests which I perform shall be reported, in the order in which they were performed, and I shall comment on any apparent anomalies. I shall use standard techniques where possible, and shall calculate coincidence probabilities based on the empirical distributions of the individual datasets. If this leads me to surprising conclusions, I shall ascertain whether this could have been caused by irregularities in the behaviour of the individual detectors.

3. For the coincidence analysis of the prototype data, I shall certainly adopt the null hypothesis, and that the amplitudes and energies etc. of gravitational waves are as set by general relativity. I shall also assume that the collapse rate in our galaxy is about one per 30 years. That is, I have a clear model which I wish to test.

4. In the case of a detection of a signal burst which is less likely, on the null hypothesis, than the chance probability of a collapse occurring during the experiment, it may seem best to suppose that a collapse has taken place and has been detected. But this will depend on the circumstances: For example, if a coincidence is of a very low probability, but the signal amplitude is so high that it cannot reasonably be explained on the basis of general relativity, then we must look again for a more likely explanation, e.g. unmodelled noise sources in the detectors. That is, I shall only postulate a candidate burst if the evidence is very strong *and* this is the best explanation for what has been observed. This assumes that such a burst falls within the accepted astrophysical model.

5. Only in the most extreme and convincing case will I postulate that a previously unmodelled source of gravitational waves has been detected. I think that LIGO and VIRGO, with sensitivities orders of magnitude better than our experiment, will have enough problems in this department.

6. We must always bear in mind that, even if we found what appeared to be a convincing gravitational wave source, the non-repeatability of the experiment, together with the unexpectedness of the detection with these prototype detectors, would mean that not everyone would believe it. Only in time, with more sensitive detectors, should we hope confidently to confirm or deny whether such a source were reasonably likely to occur during, and to be found by, the 100 Hour Experiment.

Part III

Coincidence Analysis of the 100 Hour Experiment

Chapter 5

Background to the Coincidence Analysis of the 100 Hour Experiment

This chapter will not contain much original material, being composed mostly of (i) experimental details, (ii) definitions, and (iii) a review of the analyses performed separately on the Glasgow and Garching[1] data streams. In particular, the analysis of the Glasgow data was the subject of Watkins' thesis (Watkins 1991), and I have little to add to that analysis at this point. Although the analysis of the Garching data was performed mostly by David Nicholson, no comprehensive written report on that analysis has yet been published and so it is necessary to provide some new details here. There are also points pertaining to both data sets and their analyses which have only become evident to our group since these analyses were performed, mostly due to my communications with David Robertson in Glasgow and Albrecht Rüdiger in Garching, to both of whom I am extremely grateful. These points shall be raised where relevant.

[1] Note that, throughout this analysis, I shall use the words *Garching* and *Munich* interchangeably; both refer to MPQ Garching, near Munich, Germany.

5.1 Background to the experiment

5.1.1 Objectives of the 100 Hour Experiment

No formal statement of the objectives of the joint 100 Hour Experiment has been made, to my knowledge. However, the motivations of such a long data run were recognised some time ago by all parties, and are stated in an informal report on the operation of the Garching detector during the experiment (Rüdiger 1990). They are as follows:

1. to prove that continuous operation of interferometric antennas is possible;
2. to provide long term data which can be analyzed for (non-Gaussian) noise contributions; and to histogram their frequency of occurrence versus strength;
3. to rehearse the logistics of data acquisition, data exchange and archiving;
4. and only as a faint possibility the idea of finding some gravitational waves, or some other hidden correlations between the data of distant antennas.

At the time of the experiment, in the context of the international effort, and in particular the Glasgow–Garching proposal to build long interferometric gravitational wave detectors, the first objective was seen as particularly important. Since the reliability of interferometers has been verified by the successful execution of the experiment with the high duty cycles of the Glasgow and Garching detectors, 89% and 99% respectively (and neither of which detectors were optimised for continuous trouble-free operation), the other three objectives have increased in importance. It is these three objectives, already researched in part by others, which the remainder of my thesis will address.

5.1.2 Experimental details

The 100 Hour Data Run took place between 15h 0m UT on 2nd March 1989 and 19h 0m UT on 6th March 1989. The experiment was conducted on two prototype interferometers, one at the University of Glasgow department of Physics and Astronomy, the other at the Max-Planck-Institut für Quantenoptik, Garching, near Munich. The main detector specifications are listed in Table 5.1.

Table 5.1: Specifications of the Glasgow and Garching prototypes

Detector	Glasgow	Garching
Interferometer type	Fabry-Perot cavity	delay-line
Arm length	10 m	30 m
Laser type	Ar^+	Ar^+
Position	55.86°N, 4.23°W	48.24°N, 11.68°E
Orientation of arms	193°, 283°	31°, 121°
Sampling Frequency	20 kHz	10 kHz
Absolute time accuracy	0.5 ms	$\lesssim$0.1 ms
Relative time accuracy	$< 3 \times 10^{-12}$	$< 10^{-13}$

Note:

1. Although the Garching detector has been largely dismantled at the time of writing, pending reconstruction, I shall continue to refer to it in the present tense where appropriate.

2. The *relative time accuracy* is the expected long term drift in the term $\delta t/t$; where δt is the absolute time accuracy. The Garching reference clock is at Mainflingen, (Rüdiger 1990). The Glasgow absolute reference is the M.S.F. time and frequency standard signal broadcast by the National Physical Laboratory in Rugby, used to set the absolute time on the 60 kHz laboratory reference clock. However, there was a correction of 3.2 ms in the Glasgow absolute time, due to (a) the propagation delay of the signal from Rugby, and (b) the delay in the Glasgow receiver electronics between the received time signal starting to shut down and the electronics deciding that the time mark has been made. This clock correction, supplied to me by the Glasgow group, is quoted to the accuracy shown in Table 5.1. I also refer to Robertson 1990, Robertson 1991, and Schilling 1991.

3. The misalignment of the two detectors is 5°, with both detectors having arms pointing close to the great circle joining them along the Earth's surface; and the angle between their planes is 12.3°.

4. The sampling frequency and optimum frequency of course refer to the Secondary Error Point and Secondary Feedback signals in Glasgow, and the Interferometer signal

in Garching; i.e. the gravitational wave signals. The other data recorded, e.g. seismometers etc., were generally recorded at different (lower) frequencies. Full lists of the data recorded in both detectors are given in Chapter 7. Most of these "housekeeping" data were recorded as a check on the stability and reliability of the detectors while operating, with a view to their being used to *veto* data considered unreliable.

5.1.3 Data storage and manipulation

The outputs of the detectors were written to tape in both cases. The Glasgow data were contained in 28 digital V8 mini video cassettes, each tape containing about 1.39 Gigabytes of data, while the Garching data were recorded on 94 standard 9-track 1/2-inch reel-to-reel tapes, each containing about 160 Megabytes of data. After preliminary analyses, these were sent to Cardiff for detailed analysis and comparison.

Analysis of the Glasgow tapes was performed by Watkins using a COMPAQ 386 personal computer, with a network of five Inmos transputers working in parallel. The PC was linked to two EXABYTE EXB-8200 8mm Cartridge Tape Subsystems (Exabyte tape drive, for short), one of which read the data from the Glasgow video cassettes, while the other received and wrote the results of Watkins's analysis program. His analysis is given in his thesis, where he also discusses the problems he had with this system of storage and retrieval, not the least of which were (a) the time-consuming *unpacking* of the data, which were stored on tape in a format which was easiest for the experimental team to write, but which was inordinately slow to read; and (b) that the Exabyte tape drive would sometimes reset itself during the jerky process of reading data then stopping while analysis was performed in the transputer chips. In Cardiff, the Glasgow data have recently been reduced to a format which is quicker to read, to make further analysis easier.

Analysis of the Garching tapes was performed by Nicholson on a VAX mainframe computer, which had the facility for reading the reel-to-reel tapes. The tape reader had no problems with stopping and starting during the analysis; while the amount of information stored in the Munich tapes is much less, and the unpacking was easier. Hence this analysis was quicker than the Glasgow analysis. The original Garching tapes are kept in the Department of Physics and Astronomy, U.W.C. Cardiff, although they have been copied onto video cassettes, to allow direct comparison of the two data sets on the workstations dedicated to this analysis in the Cardiff Gravitational Waves Data Analysis Group.

All of the people directly involved in the analyses of these data, the author included, have found problems with the large amounts of data involved in the experiment, which were generated in only 100 hours. These problems have sometimes been a matter of inconvenience, and, occasionally, have prevented interesting analysis. Any researchers who ignore or underestimate this problem do so at their own risk!

5.2 Individual analysis of the Glasgow and Garching data streams

In this section, I shall briefly review the analyses already performed individually on the Glasgow data and the Garching data, which is the foundation of the coincidence analysis. Before I do this, however, it is necessary to make some definitions.

Definition 5.1 *An* **event** *is a datapoint or contiguous set of datapoints in the gravitational wave stream, whether filtered or not, whose amplitudes cross a threshold.*

In the case of both the Glasgow and Garching data, the analyses done by Watkins and Nicholson use a threshold of $\pm$ 4 standard deviations from the mean, or 4σ, in both the unfiltered and filtered data streams (the mean and standard deviation were local values taken from 32768 Glasgow data points and 30000 Garching data points). Thus, *events* are relatively rare. The thresholds were arbitrary, and in no way imply the presence of a gravitational wave. In fact, the thresholds were chosen so as to permit a large enough number of events, so that we would expect a few thousand coincidences by chance, enough on which to perform some useful statistical analysis. Thresholds are related to the *false alarm rate*, i.e. the rate of occurrence of chance threshold-crossing events (see Schutz 1991).

Definition 5.2 **Experiment Time** *or* **ET** *is the time since the beginning of the 100 Hour Experiment.*

Thus an event which occurs at 16h 12m 10s UT occurs at 1h 12m 10s ET, since the experiment began at 15h 0m 0s UT. Experiment time is a useful way of referring to certain events etc. within the context of the experiment, particularly in the case of a coincidence between the detectors, since the detectors have different local times.

I shall also refer to the nth hour of data, meaning the data between experiment times nh 0m 0s and $(n+1)$h 0m 0s.

5.2.1 Individual analysis of the Glasgow data stream

Although David Robertson in his thesis performed a preliminary analysis on the Glasgow data, concerned mainly with the characteristics of the detector and data, it is necessarily brief as it was not the main topic of his thesis. The main analysis done so far was by Watkins in his PhD thesis. Those seeking anything more than the brief sketch given below should consult that work. David Robertson's thesis is also recommended for its more detailed description of the Glasgow detector and experiment details.

For the coincidence analysis, Watkins's most important contributions were a) his provision of a list of threshold crossing events and b) his analysis of the noise behaviour, both in time series and in frequency space. The event list is stored in Cardiff on video cassette, and the format of the data storage is explained in his thesis. The list is composed of events from both unfiltered and filtered data: the filtered data being the data which he match-filtered with coalescing binary templates.[2] Since the coincidence analysis which I have performed deals only with the unfltered stream, from here on I will generally talk only about the unfiltered data and results which pertain to it. In Section 6.1.4 I shall give some ideas as to how a similar coincidence analysis should be performed on the filtered event lists.

With each unfiltered time series event, most of which are only one data point long, are stored the obvious data (the time of the event, the length of the event[3], and the absolute value of the signal-to-noise of the highest S/N datapoint of the event[4]; as well as most of the relevant housekeeping data at the time of the event (seismometer signal, etc.), thus facilitating later development of criteria for removing events which are untrustworthy; either because of unusual detector behaviour or because of perturbations in the laboratory (see Chapter 7 later). The events are spread *fairly* uniformly through the experiment, although many more are seen in the first five hours or so: this is caused by a techical problem in the

[2] Nicholson has recently performed a resampling of the data at 1.25 kHz for comparison with a similar data stream in the Garching data: see Section 5.2.2.

[3] If there were no correlations between points in the time series, each threshold-crossing datapoint should be recorded separately. However, since there are too many multiple-datapoint events than could have occurred by chance, Watkins records each contiguous set of threshold-crossers as a single event: hence our earlier definition.

[4] Why Watkins took the absolute value is not clear, this making certain statistical analysis more difficult; however, since many of the longer events are oscillatory, this will not affect matters too much. In any case, one can always return to the original tapes to find the true value. Nicholson followed Watkins in recording the absolute value of the event maxima in the Garching data.

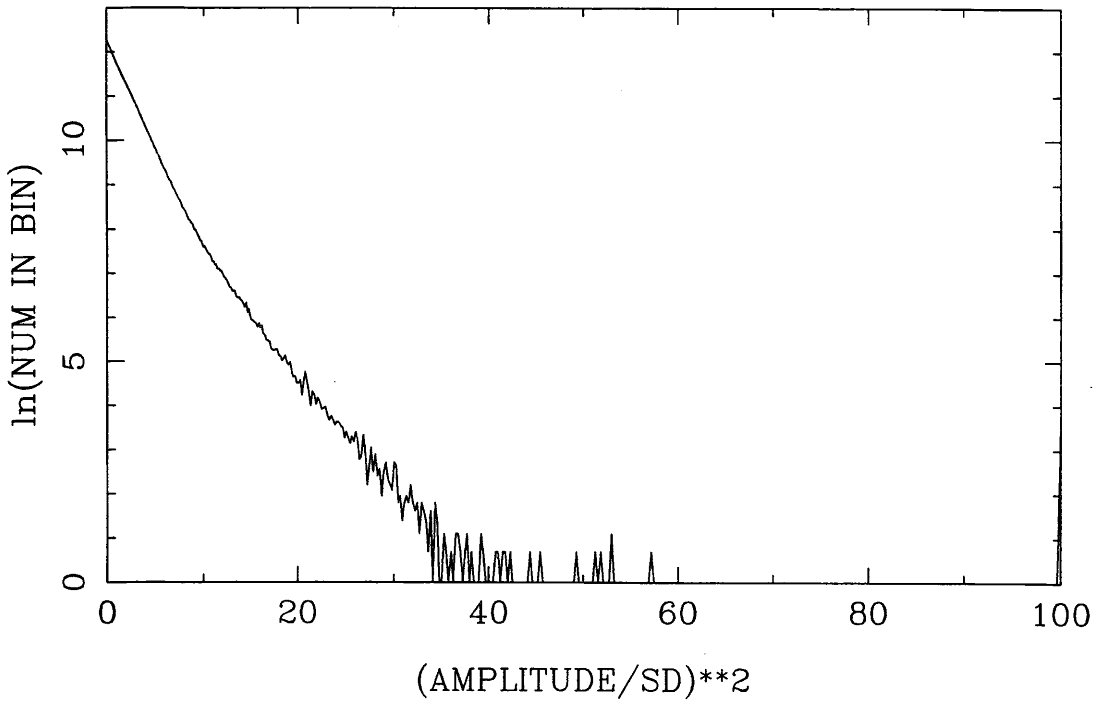


Figure 5.1: Glasgow noise distribution while the detector cavities are locked

detector, which was pointed out to me by Watkins, and it is recorded in the experiment log. At the advice of David Robertson, I have omitted these first five hours from the coincidence analysis.

Watkins also saw a clear dichotomy in the behaviour of the noise output, viz. the noise distribution of the data was close to Gaussian while the cavities of the detector were in lock (on resonance), and far removed from Gaussian when one or both of the cavities is out of lock (or in a higher mode of resonance). This is demonstrated in Figures 5.1 and 5.2, reprinted from Watkins's thesis. This gave rise to the Gaussian Parameter, devised by Watkins, as a ready reckoner as to the behaviour of the detector. The Gaussian parameter, Z tells one how close to Gaussian-ness is the noise output. We shall return to this in Chapter 7. The state of the cavities, whether or not they are in lock, is given by the Secondary Visibility signal. The Secondary Visibility signal is defined to be the intensity of the laser light reflected from the secondary cavity. See Robertson (1990).

In frequency space, the noise output of the detector was as expected, while in lock, and is shown in Fig. 5.3, again reprinted from Watkins's thesis.

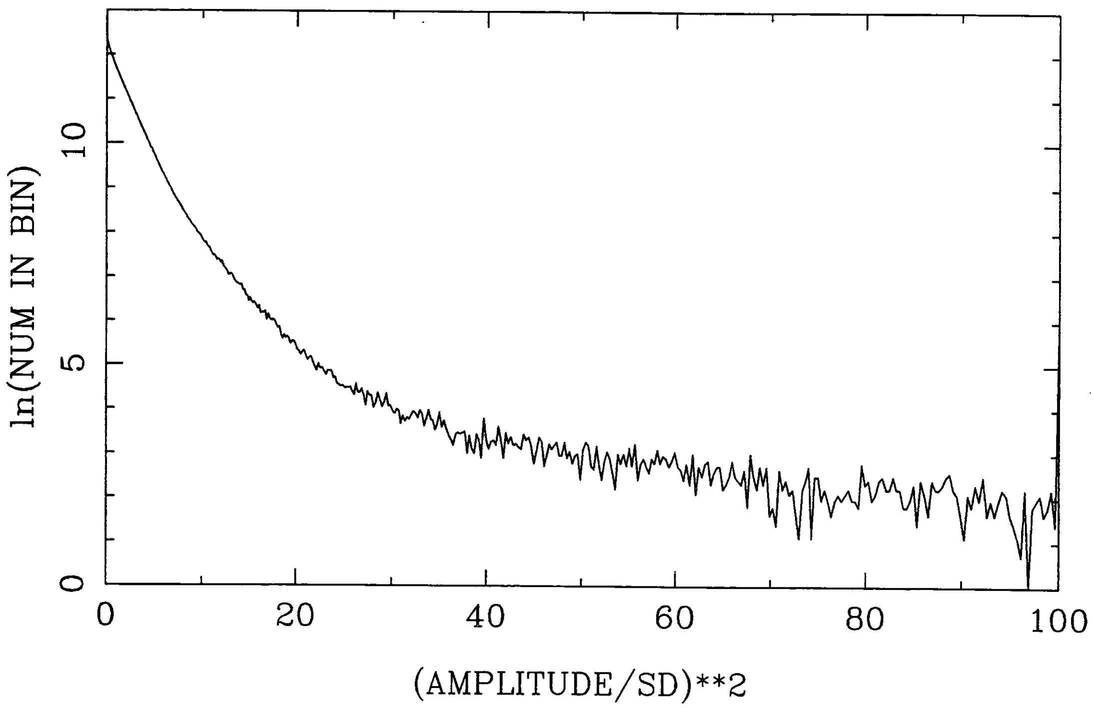


Figure 5.2: Glasgow noise distribution while the detector cavities are out of lock

5.2.2 Individual analysis of the Garching data stream

Some preliminary analysis on the Garching data was done by the German group, concerning the general behaviour of the detector, e.g. causes of loss of lock. This is reported in Rüdiger (1990). Then a paper followed, by Niebauer et al (1993), where the data were analysed in an attempt to find a continuous signal from the then candidate millisecond pulsar in the remnant of SN1987A. Of course, no signal was found; but this was an important step forward in the attacking of the problem of searching for continuous signals, both from the analytical and computational point of view.

It is David Nicholson's analysis which is most relevant to the present work, however; since again he provided a list of events, using a modified version of the code which Watkins ran on the Glasgow data. The details and results of this analysis have not yet been published, though they have been addressed at conferences, but we shall review the main points here.

Nicholson's code, like Watkins's, creates a list of events. Again there are results from unfiltered data and data filtered using coalescing binary templates. Nicholson also produces a list of events from a data stream resampled at 1.25 kHz, since we believe that any collapse

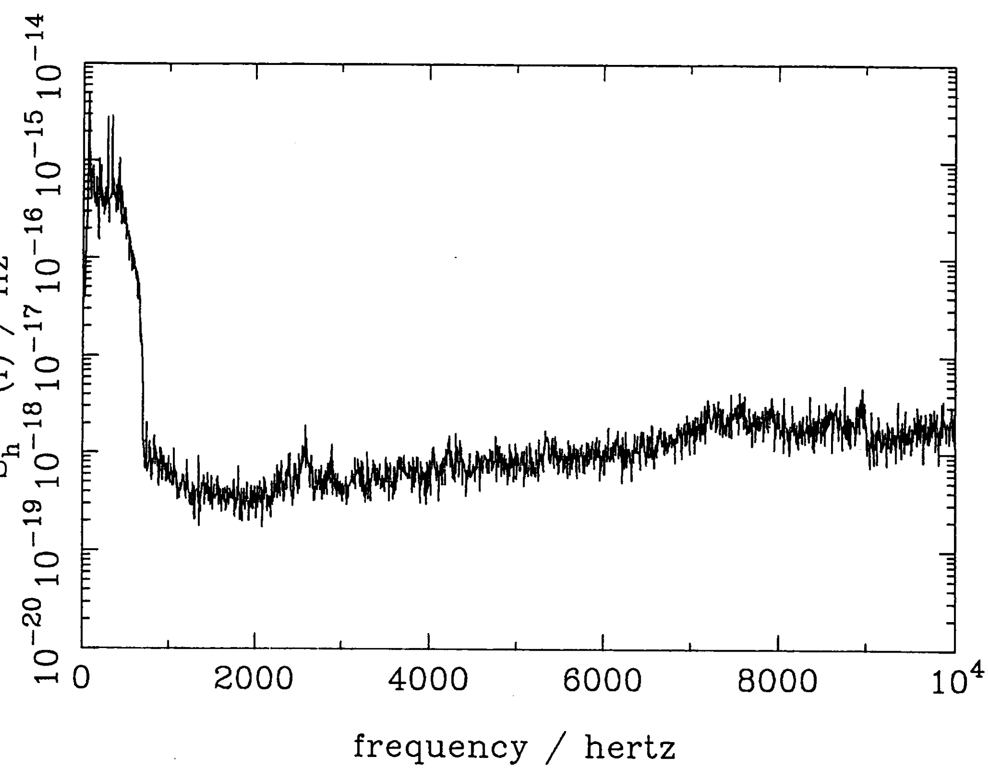


Figure 5.3: Glasgow noise spectrum

events will occur on this timescale or longer, and hence resampling would improve their signal-to-noise. Again, I shall address only the unfiltered time series results. Some of the findings are my own.

Unlike the Glasgow events, many of the events in the Garching stream are several datapoints long. It is not clear whether this is caused by filters put on the output data at the site, or whether it is caused by some internal source of correlations between points. In addition, there are more events of very high signal-to-noise in the Garching stream (see Section 8.1). Rüdiger (Rüdiger 1990; Rüdiger 1992) pointed out that there are three main known sources of such sporadic events: isolated out-of-lock periods, argon refill of the laser tube, and the laser water-cooling system. The first was noted by an alarm signal, written to the Garching data tapes, and which was later used in Nicholson's program to ignore events while the alarm is on. Although the second is noted seperately to the out of lock condition, this laser refill event did tend to throw the detector out of lock; the water cooling events did not. (In the case of the Garching detector, "out of lock" means that the servo system has lost the interference fringe at the photodiode, whereas in the case of the Fabry-Perot

Glasgow detector, it generally means that the laser resonance in the arms has been lost.) The laser refills occurred automatically in pairs of events; the pair being separated by about 126 ± 0.5 s, and there being about 200 minutes between each pair. The water cooling events occurred more erratically, again in pairs separated by about one to two minutes, but these pairs being separated by anything from five to eight minutes. I shall return to the latter two event sources in Section 7.4.2.

Nicholson also finds that the Gaussian Parameter is not such a good ready-reckoner for good and bad noise in the Garching detector as in Glasgow. However, I shall demonstrate later that there are times when the Garching Gaussian Parameter has an unacceptably high value. In Chapter 7, I shall use this and the other housekeeping data streams to attempt to veto bad data in the Garching data set.

Nicholson also found that the rms noise in the Garching detector was noticeably worse during the last 7 hours or so of the experiment. Thus I shall not consider these in the coincidence analysis later.

The noise spectrum of the Garching detector during the experiment was as follows.

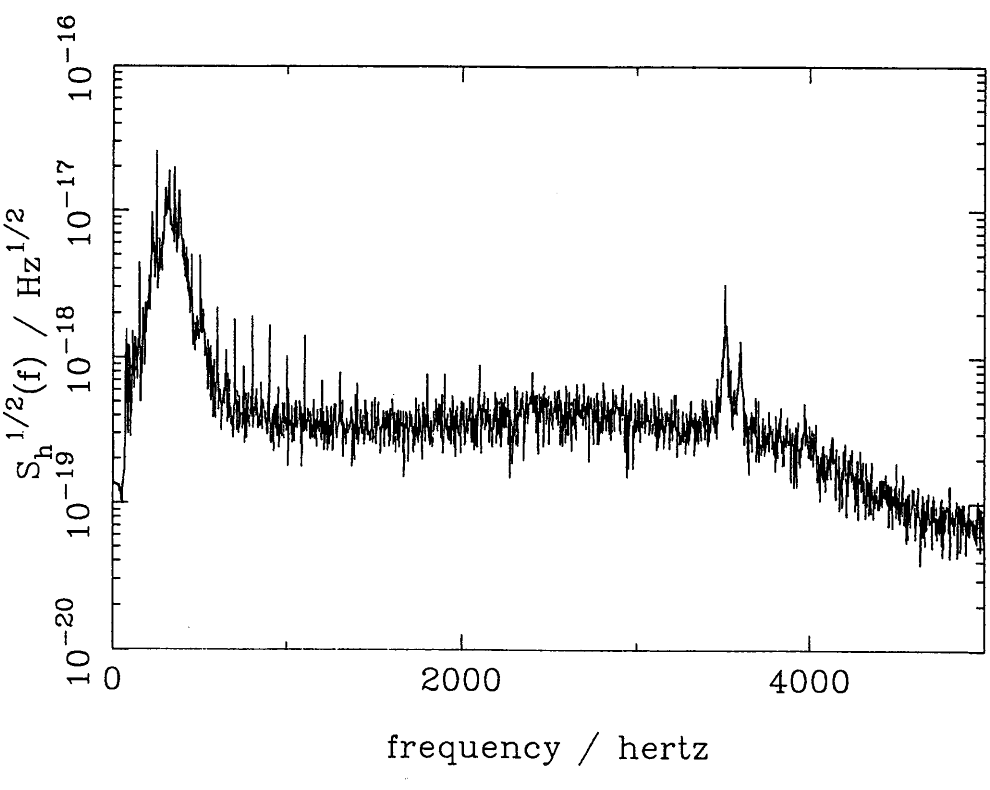


Figure 5.4: Garching noise spectrum

This concludes the review of the background and individual analysis. We shall now address the coincidence analysis of the Glasgow and Garching data.

Chapter 6

Coincidence Methodology

We shall now summarise the methodology which we shall use in the coincidence analysis of the data. Since I have already criticised what I saw as much a posteriori analysis by the RTM group (Chapter 2), I shall state my analysis methodology, as well as my expected results, before performing the data analysis. Note well that, except where otherwise stated, these methods were decided on primarily by me, after much discussion with and suggestions from Bernard Schutz and David Nicholson; and that, except where otherwise stated, the method was decided before the analysis was performed.

6.1 Analysis strategy

6.1.1 General strategy

In this analysis, I shall deal only with the time series events recorded in the two detectors. The filtered output, which addresses particular sources (data resampled at $\sim$ 1kHz primarily to search for stellar collapse, matched filtering to search for coalescing binaries) will have different analysis problems which I shall touch on (see Section 6.1.4).

The general thrust of my coincidence analysis will be a comparison of lists of threshold crossers from the two individual detector outputs. Call these lists (sequences) G and M for the event lists from the Glasgow and Munich detectors respectively. These lists will, if pure noise, in theory at least, be two independent Poisson processes with well-understood statistics. Therefore those events in coincidence between the detectors will be a random subset of the set $\mathsf{G} \times \mathsf{M}$ of all pairs of events, itself also a poisson process, except in certain circumstances which I shall explain below. If we had the empirical distributions through time and signal-to-noise space of the individual detector outputs, we could construct the expected distributions of the coincidence dataset.

We define a coincidence as follows.

Definition 6.1 *A* **coincidence** *or* **coincident event** *is a pair of* **events**, *the centroids of which have times which differ by less than or equal to the light travel time between detectors; in this case 4.6 ms.*

By the *centroid* of an event, I mean the midpoint in time of the contiguous set of datapoints which cross the threshold. Clearly, if the event is in fact only one data point long, then its centroid is the time of that singular threshold crossing datapoint.

The reader may ask: Why use the centroids rather than either (a) the peak values of signal to noise or (b) accept an overlap of any two points in the events. The answer to the first suggestion lies in the potentially ruinous unpredictability of the noise, as the following example demonstrates. Suppose for the sake of simplicity that the two detectors are operating at the same sensitivity, and that a gravitational wave passes the detectors at a moderate signal to noise of, say, 6. Then the detector noise will add onto the observed signals linearly and in an unpredictable way. The following situation may arise:

If the gravitational wave signal were coming in at an angle such that the time delay between observations is, say 4.2 ms, then an unfortunate asymmetry in the noise contributions of the detectors can produce the situation shown in Fig. 6.1. This would not normally matter, but when the events are close to the edge of the maximum wave travel time (4.6 ms), this asymmetry could be enough to push the peak values outside this time window, and this would not be regarded as a coincident event by an analysis system based on comparing peak values. One could argue that this set of circumstances would depend on two improbable events, viz. (1) the coincidence occurs close to the edge of the coincidence window and (2) the noise adds to the signal in this unfortunate way, making two peaks at opposite ends of the events; but nonetheless, this situation will inevitably occur from time to time in working observatories.

As for method (b), if we allow any kind of overlap between events, modulo the wave travel time window, we will not only allow all real signals to be in coincidence, we will allow many which clearly cannot be real signals. A real signal will arrive at the same time in

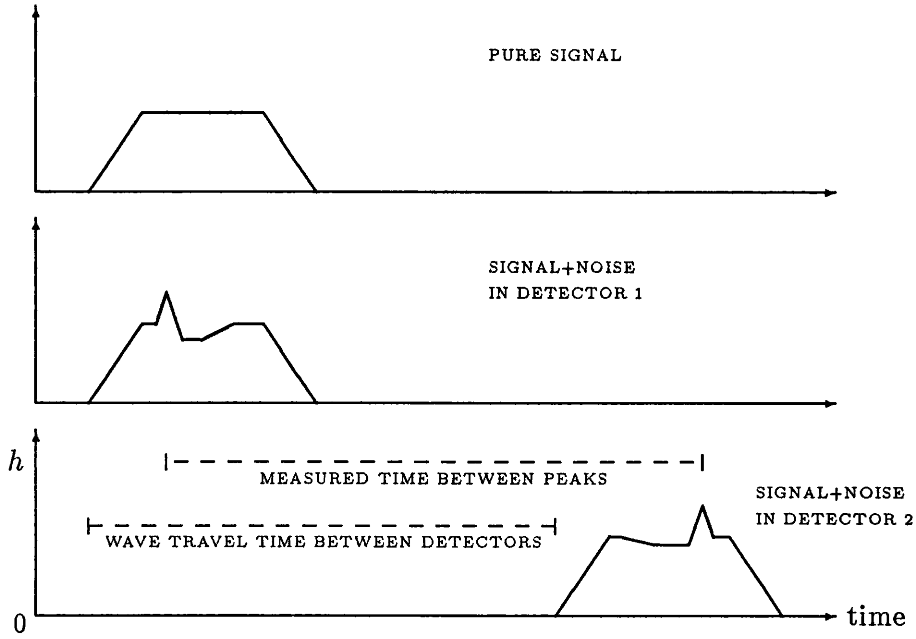


Figure 6.1: Example gravitational wave signal in two detectors with independent noise. Although the same real signal is present in both outputs, the peak values of signal-to-noise output here are more than 4.6 ms (the maximum wave travel time) apart.

both detectors (not necessarily above threshold, since the threshold depends on the detector sensitivity), within the maximum light travel time, and will end at the same time, with the same time delay as in the start of the event. Thus, for example: if an event started in Detector 2, 4.2 ms after ending in Detector 1 and the event is 1 ms long, then method (b) would describe this as a coincidence; but taking the simple event morphology above, neither the starts nor the ends nor the centroids of the two events would be within 4.6 ms of each other (in fact, all would be about 5.2 ms apart). So this would not be reasonably described as a real source, and should not be defined as a coincidence.

It has to be said that one could contrive pathological event morphologies and detector sensitivities such that even the centroid method would not be optimal. Since our model of sources is probably far from complete or definitive, there is little we can do about this, with two detectors with mismatched sensitivities as the Glasgow and Garching detectors sometimes are. In any case, our definition shouldn't make much difference to the coincidences recorded since most events are much shorter ($\sim$ 5 datapoints=5×10^{-4} s) than the coincidence window ($\pm4.6 \times 10^{-3}$ s). It is really the difference of these methods *in principle* which is worth stating explicitly, for the interest of future working observatories and data analysis systems.

We shall not use summation of the gravitational wave signal or a cross-correlation of the two data streams. Although both could give reasonable results in principle, since the detectors are closely aligned, these are large tasks in themselves and best left to another work.

Hence, using this definition of "coincidence" given above, looking at the whole 100 hours of data, we shall obtain a *coincidence list* of those pairs of events which pass threshold in both detectors and whose centroids occur within ± 4.6 ms of each other. Again, any coincidences which we may find are not necessarily real gravitational wave sources, and we must adopt the null hypothesis a priori, as suggested by our current model, which is that all the coincidences are random noise events (see also Section 6.2).

Now, we would like this list to be a Poisson process with no sequential dependence between events, as we suppose the Glasgow and Garching event lists to be. This would greatly simplify the statistics. The list will be a Poisson process if the events in each of the individual lists are much further apart than the coincidence time window, as demonstrated crudely in Fig. 6.2.

In Fig. 6.2, I show two cases. In the first, Fig. 6.2(a), the event centroids (represented by small vertical strokes) are typically much further apart than the length of the light window; in Fig. 6.2(b), they are typically slightly closer together than the light window. The corresponding coincidence lists are also shown.

Note that in Fig. 6.2(b), it sometimes happens that the same event in Detector 1 is in coincidence with more than one event in Detector 2, and vice-versa. These "multiple coincidences" are clearly not sequentially independent; and since sequential independence (the *zero memory property*) is an important property of Poisson processes[1], the coincidence list of (b) is not a Poisson process. In fact, there will probably be some multiple coincidences between Poisson processes with *any* mean time between events, because a few will probably be close together by chance. But certainly, the problem becomes worse as the mean time between events in either data stream becomes shorter.

In fact, in the 100 Hour Experiment, during the hours of data we are using in this analysis

[1]It is a well-known result that Poisson processes (i.e. those processes where the arrival times of events are distributed uniformly) are the *only* continuous processes with this property.

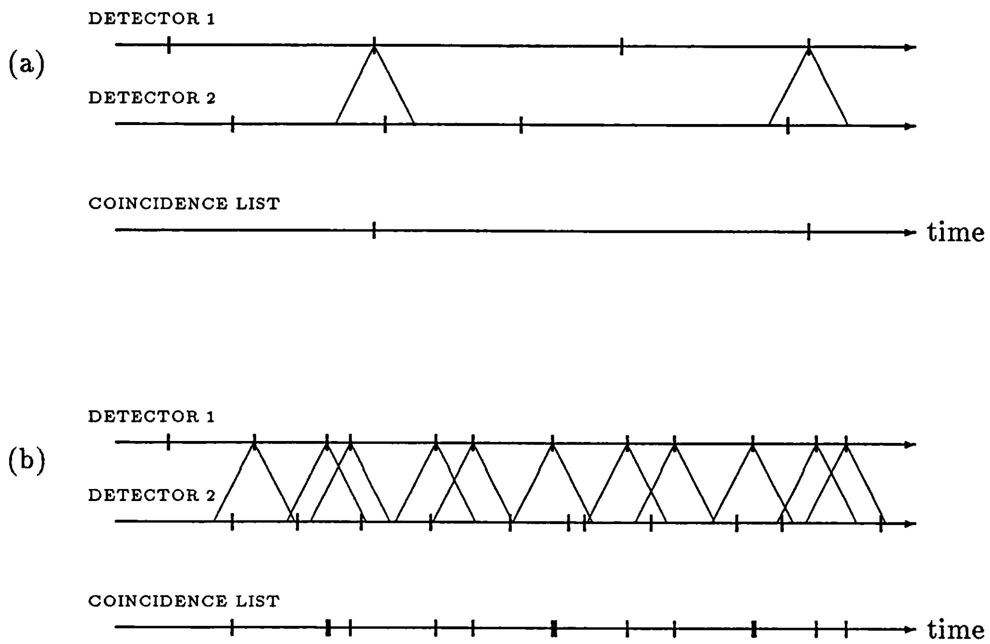


Figure 6.2: Coincidences produced by individual event lists in two cases. In (a), the average time between events in the individual lists is much longer than the length of the coincidence window; while in (b), the average time between events is of the order of, or shorter than, the length of the coincidence window.

(5h–93h E.T.), the detectors record about 2670000 (Glasgow) and 1340000 (Munich) events (before vetoing), which is an average separation of 0.12 s per event (Glasgow) and 0.24 s per event (Munich), compared to the much shorter 9.2×10^{-3} ms size of the acceptance window. So, provided the event streams are reasonably uniformly-spaced in time, the coincidence list will to all intents and purposes be a Poisson process. We can also, therefore quickly calculate the expected number of coincidences, N_c, which is

$$N_c = \frac{N_G N_M \times 9.2 \times 10^{-3}}{l_s}, \tag{6.1}$$

where N_G and N_M are the numbers of Glasgow and Garching events, l_s is the length of the experiment in seconds, and the 9.2×10^{-3} comes from the coincidence window being ± 4.6 ms. Taking $l_s \sim 88$ hours we get

$$N_c \sim 1.04 \times 10^5. \tag{6.2}$$

6.1.2 Assessment of significance of results

The two most important numbers associated with each coincidence shall be (a) its inferred gravitational wave strain amplitude, h, and (b) the probability of its chance occurrence during the experiment.

The strain amplitude should be simply calculated by multiplying the observed signal-to-noise ratio in each detector by the sensitivity of the detector at the time. The two inferred h values would almost certainly differ, even if the coincidence were real, due to different and independent noise contributions in the two detectors. If one wishes to infer the most likely original source amplitude, given the hypothesis that a real wave caused the detection, one would need to make some sort of average of the observations, weighted by the detector sensitivities, and weighted also by the expected model of observable sources. In our analysis, a simple mean of the two should suffice. This will give us the upper limit on broadband burst sources for the experiment.

The probability of occurrence is more tricky. Certainly, the large number of high amplitude events would render absurd a calculation based purely on the assumption of a Gaussian noise distribution in both detectors, so some sort of empirical calculation is required. The two most obvious are:

1. to plot the distributions of the signal to noise of the individual detectors, associate a probability (based on the frequency of occurrence) of each S/N value, and multiply these empirical probabilities together to get the probability of a given coincidence;

2. to perform the coincidence analysis for other unphysical windows, having the same width but offsetting the data streams by times larger than the light travel time between detectors, ensuring that these offset datasets have no common gravitational waves, and using these as control sets against which to compare the physical time window. For example, if we were to generate enough control datasets, say 1000, and a coincidence in the physical dataset were the largest out of all 1001 datasets, then we would ascribe a probability of chance occurrence per experiment of $\lesssim 10^{-3}$.

The main problems with method 2 are interpretation and programming inexpediency. By interpretation I mean, what does one mean by "the largest coincidence in the ensemble of datasets", when one is dealing with a list of *pairs* of values from two detectors? Programming inexpediency breaks down into programming time and error, running time of program, and storage and collation of results. Both methods, of course, should give the same answer, so we shall use the simpler method 1.

Should we find what I shall loosely call a candidate real event (i.e. a coincidence with very low probability of occurrence by chance in noise), I should address its physical origin in a simple and clear way, as follows. Given a coincidence, its probability of occurrence in noise during the experiment can be calculated empirically. We also have a model of burst sources in our galaxy which surmises that there will be one observable collapse about every thirty years on average. This means that the probability of such an event occurring during the 100 hour experiment is 4×10^{-4}. Hence, we must treat very seriously any coincidence which has a noise probability of less than, or of the order of, this number. This would be subject to there being nothing else in the experimental results which suggest our model is wrong in which case we need first to re-examine the model. Examples of this would be if the event in question were much too strong or too short or too long, or if there were several or even dozens of such events.

6.1.3 Vetoing poor data

It is known that the operators of the two detectors also recorded *housekeeping data*, that is, data other than detector gravitational wave output that were recorded in the laboratory at the time of the experiment, to indicate the behaviour of the detector, environmental effects, etc. These shall be used to remove or *veto* gravitational wave data which for one reason or another are not trustworthy. We shall deal with this in full in Chapter 7. Clearly, this will affect the empirical calculation of the probability of coincidences. Thus, if we veto some of the coincidences, we must also apply the same vetoes to the individual data streams to obtain the correct distributions from which the coincidences were drawn.

6.1.4 Brief preview of analysis of filtered output

As I have already said, I shall not attempt to analyse any of the filtered data sets. For future such analysis, I make here some suggestions and guesses as to how analyses of these data sets should proceed.

The analysis of data low pass filtered at 1.25 kHz would, in principle, be very similar to the broadband search performed here. The calculation of the probability of each coincidence should, of course, be based on the individual *resampled* datasets. The vetoes I use for the broadband search (see Chapter 7) would change somewhat. The *group* vetoes (q.v.) would be the same , as the group housekeeping values are constant over timescales of $\sim$ 1 s. The *event* vetoes (q.v.), on the other hand could change, since some of the event housekeeping data are recorded at a rate higher than 1.25 kHz. In this case, one could either take the nearest housekeeping record, or the worst in the group, or even resample the housekeeping data at the same rate as the gravitational wave output. Apart from these things, I believe the analysis of resampled data would in principle be the same as that for broadband data.

More problematic is the analysis of matched filter events. The first problem is computational, by which I mean the coincidence program is greatly complicated by the inclusion of algorithms to find coincidences between matched filter events as well as time series events. I have solved this problem in the case of sequential programming, checking each filter in a group of data sequentially (see Section 6.4). However, I strongly suspect that this could be simplified and speeded up by using parallel programming, with one or more sets of filter events being dealt with simultaneously by each processor of a parallel network, the way they were analysed by Watkins's code (Watkins 1991). This process could be extended to cover the coincidence analysis of all filters used in the individual analysis, in our case the only other being the data resampled at 1.25 kHz.

The other main problem with matched filtering is the fact that the outputs of each filter are not completely independent of each other. In some very brief coincidence analysis I performed on stretches of matched filter events, I found that often the same stretch of data could produce output which was above threshold not only in more than one matched filter, but in the broadband time series data as well. This is contributed to by the relatively low threshold we use, but in principle it could happen for any threshold, if we chose two filters which were close enough in mass parameter and phase. This raises two questions:

1. Should we count as coincidences those filter events which arise from two different but neighbouring filters in the two data streams? This is particularly important when the filters of the *same* mass parameter in each data stream do *not* both detect the event. Whether this happens will depend partly on how far apart in mass parameter space the filters are placed, and partly on the threshold. This question is probably best treated analytically, and is outside the scope of this thesis.

2. Given that one would have many lists of coincidences, each corresponding to one filter,

with each coincidence having a corresponding probability of occurrence, it must be noted that the lists of probabilities may not be independent of each other; e.g., a coincidence in one filter may make it more likely that there will be a coincidence in neighbouring filters. This could affect a calculation of the probability of occurrence of interesting events. Again, this would be less likely to occur for higher thresholds and/or more widely spaced mass parameters, but it must be addressed theoretically at least. Again, this is outside the scope of this work.

These questions are really the same; the first takes the point of view of a signal being detected by two or more neighbouring filters, the second addresses the behaviour of the noise. In both cases it is the signal-to-noise, and hence the threshold, which will be affected in a way which is not obvious. The question of how closely spaced a suite of filters should be has been addressed elsewhere (e.g. Sathyaprakash & Dhurandhar 1991) from a statistical point of view, but at present, the actual number of filters required is being debated (see e.g. Cutler *et al.* 1993).

6.2 Astrophysical model

I shall now quickly summarise the relevant aspects of the astrophysical model, from which any real events could come. It must be stressed again at this point that this is only a model, and is there to be tested, shot down, or strengthened by other researchers or by the results of this or any future work in the subject. In particular, I ask the reader to remember the huge differences between prior model and experimental results already seen in other young branches of astrophysics, such as the microwave background, solar radio emission, the solar neutrino problem, unpredicted sources such as pulsars, QSO's, etc. Although we certainly don't expect to see any real gravitational waves at our sensitivity level, we may not find what we expect.

6.2.1 Expected sources

The main expected sources of observable gravitational radiation are well-known to those in the field (see, e.g., Thorne 1987; Schutz 1988b; Nicholson *et al.* 1992). They are as follows:

1. stellar collapses

2. coalescing compact binaries

3. continuous wave sources, such as rotating neutron stars

4. the stochastic background due to big bang echo, cosmic strings, etc.

The analysis covered in this thesis can deal only with (1), but, by default, could unearth any similar strong short-duration unmodelled events which may arise.

The waveform from a core collapse has been simulated numerically, but only in the case of axisymmetric collapse, which we do not expect will give much, if any, gravitational radiation (see Thorne 1987). In the case of non-axisymmetric collapse, we have only a very vague idea of the expected waveform. The collapse timescale for a 1 $M_\odot$ star is around 10^{-3}–10^{-2} s, while the natural period of vibration and the light-crossing time of a 1 $M_\odot$ black hole are both around 10^{-4} s. Thus, it may be that the right frequency at which to look for such sources is around 10 kHz, i.e. the frequency of the data taken during the experiment. We shall concentrate on this scenario. Alternatively, it may be that collapse sources would be better seen at around 10^2–10^3 Hz, and work on these data, resampled at 1.25 kHz, is going on simultaneously with the writing of this thesis (Hough *et al.* 1993).

6.2.2 Effects of detector geometry

The relative orientation and separation of the two detectors has a bearing on the coincidence analysis. Firstly, I must point out that the two detectors are closely aligned, i.e. one is rotated by only about 5°with respect to the other (modulo interferometer rotational symmetry) when shifted along the great circle joining their locations on the Earth. Also, the planes in which they sit intersect each other at an angle of 12°. Thus for sources at most points on the sky, the detector response will not be greatly affected by relative detector orientation.

Below, I shall show that for sources distributed uniformly on the sky, there is no preferred delay between detections, i.e. that the distribution of expected detection time delays is flat. Thus I shall conclude that we do not need to weight the amplitude or probability of observed coincidences according to the time delay between the two events of a coincidence.

Consider the two detectors, G and M, placed on the baseline joining them, as in Fig. 6.3 (a). Let a source be denoted S. Then directions GPS and MS are parallel, and the path difference is l.

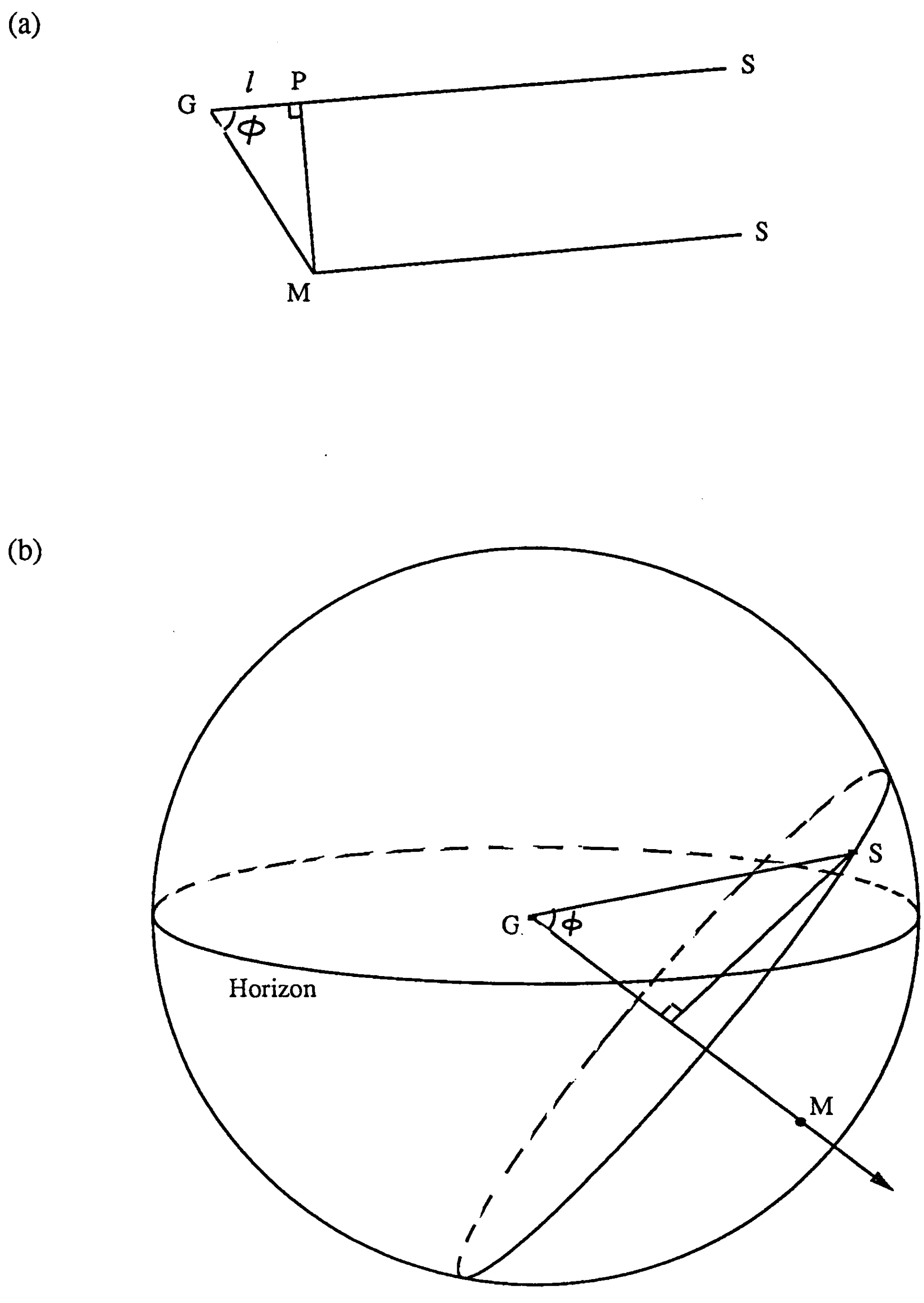


Figure 6.3: Relative geometry of two detectors, G and M, observing a distant source, S

Denote angle $\widehat{\mathrm{SGM}}$ by ϕ. Then we have

$$\cos\phi = \frac{l}{GM}, \tag{6.3}$$

and

$$\tau = \frac{l}{c}, \tag{6.4}$$

where τ is the time delay between the reception of the signal by M and by G, and c is the speed of light.

Now project onto G's celestial sphere as in Fig. 6.3 (b). ϕ is still the angle $\widehat{\mathrm{SGM}}$. Define a spherical coordinate system centered on G, with a pole at M (without loss of generality). Then ϕ is the colatitude of S in this system. Also, define a longitude coordinate, λ, with respect to an arbitrary meridian $\lambda = 0$.

Rotate SGM about GM to form a small circle, ϕ = constant, the locus of point sources on the sky with the same time delay, τ, between detections. This circle is great (maximal) for $\phi = \pi/2$ and pointlike for $\phi = 0$ and $\phi = \pi$.

Now, for sources distributed randomly and uniformly on G's celestial sphere, the probability of a source being in an area delimited by colatitudes $\phi, \phi + \Delta\phi$ and longitudes $\lambda, \lambda + \Delta\lambda$, for $\Delta\phi$, $\Delta\lambda$ small, is the area of this small region divided by the area of the sphere, i.e.

$$\mathrm{P}(\phi, \phi + \Delta\phi; \lambda, \lambda + \Delta\lambda) = \frac{\Delta\lambda \sin\phi . \Delta\phi}{4\pi}. \tag{6.5}$$

Then the probability of having colatitude between $\phi, \phi + \Delta\phi$ irrespective of longitude is obtained by letting $\Delta\lambda \to 0$ and integrating with respect to λ from 0 to 2π; i.e.

$$\mathrm{P}(\phi, \phi + \Delta\phi) = \int_0^{2\pi} \frac{\sin\phi\Delta\phi}{4\pi} d\lambda \tag{6.6}$$

$$= \frac{\sin\phi . \Delta\phi}{2}. \tag{6.7}$$

Then the probability of having colatitude value between ϕ_1 and ϕ_2, irrespective of longitude, is obtained by letting $\Delta\phi \to 0$ and integrating with respect to ϕ from ϕ_1 to ϕ_2; i.e.,

$$\mathrm{P}(\phi_1, \phi_2) = \int_{\phi_1}^{\phi_2} \frac{\sin\phi}{2} d\phi \tag{6.8}$$

$$= \frac{1}{2}(\cos\phi_1 - \cos\phi_2). \tag{6.9}$$

By definition, the integrand of Eq. 6.8 is the marginal probability density function of ϕ, i.e.

$$\mathrm{p}(\phi) = \frac{1}{2}\sin\phi, \qquad (0 \le \phi \le \pi). \tag{6.10}$$

Clearly, the probability density of events is higher around $\phi = \pi/2$, and so more events will take place around this colatitude than at the poles $\phi = 0$ and $\phi = \pi$.

Now, we are really interested in the marginal probability density function, $\mathrm{p}'(\tau)$, of the time delay between the two observations of a coincidence, and the corresponding cumulative distribution function, $\mathrm{P}'(\tau)$. By definition,

$$\begin{aligned}
\mathrm{P}'(a) &= \text{Probability}(\tau \le a) && (6.11)\\
&= \text{Probability}\left(\frac{\mathrm{GM}}{c}\cos\phi \le a\right) \qquad \text{by 6.3, 6.4} && (6.12)\\
&= \text{Probability}\left(\cos\phi \le \frac{c.a}{\mathrm{GM}}\right); && (6.13)
\end{aligned}$$

and since cosine is bijective and monotonic decreasing on $[0,\pi]$ we have

$$\begin{aligned}
\mathrm{P}'(a) &= \text{Probability}\left[\phi \ge \cos^{-1}\left(\frac{c.a}{\mathrm{GM}}\right)\right] && (6.14)\\
&= \int_{\cos^{-1}\left(\frac{c.a}{\mathrm{GM}}\right)}^{\pi} \frac{\sin\phi}{2}d\phi \qquad \text{by 6.10} && (6.15)\\
&= \frac{1}{2}\left(\frac{c.a}{\mathrm{GM}} + 1\right). && (6.16)
\end{aligned}$$

Hence, by definition,

$$\begin{aligned}
\mathrm{p}'(a) &= d/da\left[\frac{1}{2}\left(\frac{c.a}{\mathrm{GM}} + 1\right)\right] && (6.17)\\
&= \frac{c}{2\mathrm{GM}}; && (6.18)
\end{aligned}$$

i.e.,

$$\mathrm{p}'(\tau) = \frac{c}{2\mathrm{GM}}. \tag{6.19}$$

That is, the probability density function of time delays between the two detectors is a constant for $-\frac{\mathrm{GM}}{c} \le \tau \le \frac{\mathrm{GM}}{c}$ and 0 elsewhere. (Note that, as expected, $\int_{\phi=0}^{\pi} \mathrm{p}'(\tau)d\tau = \int_{-\frac{\mathrm{GM}}{c}}^{\frac{\mathrm{GM}}{c}} \frac{c}{2\mathrm{GM}}d\tau = 1$.) Consequently, one is no more likely to obtain coincidences at one delay than at any other, within the limits $-\frac{\mathrm{GM}}{c}, \frac{\mathrm{GM}}{c}$ ∎

6.3 Expected laboratory-based effects

I shall briefly review here the effects which were known to have occurred "on the ground" during the experiment, which were not directly or intentionally measured by the housekeeping data, and which may affect the analysis. These include:

- laser water cooling events (Munich),
- laser refill events (Munich),
- A/D sawtooth in the Table Seismometer data (Munich),
- low-pass-filter-induced "sinc glitch" events (Munich),
- data repetition on tape (Glasgow), and
- "dwangies" (Glasgow and Munich).

The Munich laser was the source of two sets of spurious sporadic events. To prevent the laser from overheating, it was fitted with a water cooling system which activated twice every seven minutes or so in a fairly regular pattern, though this pattern was not regular enough to be useful (see Section 7.4.2). The switching of the valve caused an oscillatory event to be recorded, of maximum amplitude $\sim$ 10 standard deviations and duration $\sim$ 1 ms. These events did not cause the detector to lose lock, and so are difficult to remove. Perhaps worse, their oscillating above and below threshold caused them to be recorded as many separate events in some cases. Hopefully, this will not happen in future such experiments, since the temperature threshold on the thermostat can be adjusted so that this switching rarely happens (Rüdiger 1992). As a point of interest, these signals are coupled in at the leads connecting the detector to the data acquisition computer; so they could perhaps have been removed quite easily using a mains pulse detector as in the Glasgow laboratory (see Chapter 7).

The Munich laser also refilled itself with argon, and was recorded having done so 24 times during the experiment, each time causing a pair of audible clicks separated by (126 ± 0.5) s and causing loss of lock. This would have activated the alarm, and so any events thus caused would have been vetoed by Nicholson's event program.

During the experiment, a short sawtooth (of 1 ms duration) was added at the beginning of every minute, as a check that no extra pulses were seen by the A/D converter's trigger input (Rüdiger 1990). This then appeared very strongly in the Munich Table Seismometer stream (see later).

Some very strong oscillatory signals were seen by Nicholson in the Munich gravitational wave data, always in the first 10 datapoints or so of the group (and also at the end of each group, which were overwritten since Nicholson's analysis of the Munich data was like Watkins's, in that succeeding groups overlapped in time). He believes that these were sinc-type signals induced by the Munich Bessel low pass filter. The coincidence program removed these events "on line".

It was found that twice in the Glasgow data, it happens that there is a block of data which is recorded many times ($\sim$ 50 times), giving extra events in the event list and later in the coincidence list. The signature of this is a repeated *block time* in the data. It is also seen as an unnaturally high bin or bins in some of the housekeeping streams (see Chapter 7), the latter indicating that the same housekeeping value from the same event has been recorded many times. This is easily warded against with an extra line in the coincidence program to remove the repeated blocks, or by finding then removing the culprits in the event list after running the coincidence program.

Finally, we come to what are becoming known as "dwangies" — this is a general name for sporadic, short, large amplitude, oscillatory events seen in the gravitational wave output of the Glasgow and Garching detectors, and also seen in the output of the Caltech prototype (Zucker 1992) where they are referred to with the less colourful title "glitch". These are thought to be the result of some kind of perturbation affecting the detector or its electronics, followed by attempts by the servos to retain lock, resulting in oscillatory output. These dwangies are sometimes, but not usually, associated with loss of lock in our experiment. In the Munich output, the only dwangie-like events of which we are aware are those caused by the water-cooler and laser-refill events. In the case of Glasgow, it is not so clear what causes them, but they do occur. It is hoped that the Glasgow dwangies will show themselves in data streams other than just the Secondary Error Point, e.g. the Primary Error Point, and that they will thus be vetoed.

Of course, such things as loss of laser cavity resonance, seismic perturbations, etc. will be dealt with more explicitly in Chapter 7.

6.4 Coincidence program

I shall now briefly discuss the program which I wrote for the coincidence search, listed in Appendix A, and then point out a few changes made by Nicholson before the program was run for a second time.

6.4.1 My coincidence program

The main task of the program originally written was to search through two lists of events (the "event list"), where the data given had a particular group structure involving the threshold crossing events output from many filters; and make new lists of pairs of events (the "coincidence list"), one event of a pair from each dataset, where the two events of the pair were less than or equal to 4.6 ms apart in time. The program is listed fully in Appendix A. The program and the following descriptive outline will only make sense to those acquainted with Watkins's program and the data output format thereof (Watkins 1991).

The main program TESTCC had the basic task of comparing each event datapoint in turn[2] for each filter, to check for coincidences, and writing any coincidences to an output file appropriate to the current filter.

The data were read in using subprograms TESTLGB and TESTLMB for Glasgow and Munich data respectively.

Because of the block structure of the Glasgow and Munich event lists (see Watkins 1991), where the events are recorded chronologically for the time series (TS) events, then chronologically for all matched filter events of mass parameters around 1.4 $M_{\odot}$ (1.4 solar masses), then chronologically for all matched filter events of mass parameters 2-6 $M_{\odot}$, it was necessary to tabulate each group of events in an array, where each column corresponds to one filter only, and the events in each column are chronological. This is done by subprograms GTABLE and MTABLE. This enables the main program to check each filter, and then to run through the events chronologically in each filter.

Again because of the complicated data structure of the event lists, it was necessary to have two subprograms which could quickly find the records corresponding to the events currently being examined in the main program. These subprograms are GEVLOC and MEVLOC.

[2]in this version of the program, datapoints are treated independently irrespective of whether they form a contiguous set of threshold crossers.

To summarise, the program will do the following:

- load in a block of Glasgow (Munich) data as required, using subprogram TESTLGB (TESTLMB);
- copy the next group of Glasgow (Munich) data into the array GGROUP (MGROUP);
- run subroutine GTABLE (MTABLE) to sort the events by filter;
- for each filter, loop through the events of each group, keeping the Glasgow and Munich events synchronised, and skipping back through the Munich events when a multiple coincidence is encountered (back to the "trigger event", the first Munich event to be in coincidence with the current Glasgow event);
- locate within the current GGROUP (MGROUP) the time of the current Glasgow (Munich) event (using subprogram GEVLOC (MEVLOC));
- if the events are in coincidence, within the light travel time, write the pair of events, including all the commensurate data of interest (time, S/N, housekeeping, etc.) to the output file appropriate for this filter;
- new Glasgow (Munich) event as required;
- next filter for these groups of data, as required;
- new group of Glasgow (Munich) data copied into array GGROUP (MGROUP) as required;
- new block of Glasgow (Munich) data read from disk, as required.

Note that, in order to make sure that each Glasgow event of the current Glasgow group was synchronised with the Munich data, it was necessary to always keep *two* Munich groups in memory, viz. those Munich groups "surrounding" the current Glasgow group in time.

6.4.2 Nicholson's adaptation of the coincidence program

After extensively testing the above program and running it to provide large samples of coincidence lists, it was decided that some modifications were desirable. The main motivations behind this were firstly, to make my program easier to read and digest by scientists not in our group in Cardiff and, secondly, to include lists of coincidences corresponding to unphysical delays (see Section 6.1.1). For reasons of running time and disk space, only 1000 artificial delays could be used. The main modifications were as follows:

- unphysical delays were included for comparison with physical delay;
- either of the first Glasgow or Munich groups could be the earlier, contrasting to my program which would not achieve synchronicity if Munich were earlier;
- only the TS (time series or delta function filter) events were used;
- contiguous sets of threshold crossers were treated as a unit, with one characteristic time (as discussed in Section 6.1.1), since, particularly in the case of Garching, sequential datapoint threshold crossers were found to be not independent;
- the window was widened to ± 5.0 ms to produce Fig. 8.2, with unphysical events ($4.6 \leq |\tau| \leq 5.0$ ms) to be removed later;
- on line vetoes were included to remove those coincidences where:
    1. the Munich Alarm was on,
    2. the Glasgow calibration flag (see Watkins 1991) was on, or
    3. the Munich event occurred in the first 10 datapoints of the Munich group (characteristic of the short sawtooth in the A/D converter: see Section 6.3).

We now address in more detail the question of removing or vetoing untrustworthy data, and this is the subject of the next chapter.

Chapter 7

Vetoes

As already indicated, there are times during the experiment when the output of either or both detectors cannot be trusted. This is when, for example, a detector was out of lock, or there was significant seismic disturbance of a detector, or there was audible noise in the laboratory. Since any gravitational wave excitations occurring at such a time would always be treated with suspicion, it is better that the data recorded at such times be removed, or *vetoed*.

The slight disadvantage of this procedure is that one throws away data which may contain real gravitational waves. However, as shall be seen, the actual fraction of data so vetoed is not great compared to the full 100 hours of the experiment, and so the chance of accidentally removing a real gravitational wave signal would be very small. On the other hand, the advantages of this vetoing procedure are clear: by removing events caused by effects outside the detector, one would hope to make the noise distribution more Gaussian, and hence easier to model and analyse. It would also, potentially, have the effect of preferentially removing long structured events (~ 1 ms, rather than the more typical noise generated events of length one or two data points, ~ 0.1 ms) caused by external perturbations of the detector, which could look much like a real gravitational wave is expected to look. As a result of this, the false alarm rate would be lower for a given threshold, one could set better limits on h for the experiment, and, if there were any real gravitational waves detected, their signals would be more outstanding.

In this chapter, we shall look at the housekeeping data recorded by the experimenters, and we shall assess which ones may be used as vetoes against untrustworthy data. We shall outline the method for this assessment before proceeding; and we shall choose, in as unbiased a way as possible, the values of the housekeeping data which indicate untrustworthy data.

7.1 Data available for vetoing

We begin this chapter, then, with a list of all the data recorded by the experimenters in Glasgow and Munich, indicating whether they were retained by Watkins's and Nicholson's individual analysis programs, and an indication as to whether these data bear further investigation for the purpose of being used as vetoes. The list of Glasgow data (Table 7.1) was taken from Robertson (1990), and the list of Munich data (Table 7.2) was taken from the Munich documentation on the experiment (Rüdiger 1990). Both these references contain detailed explanations of the signals and how they were taken. The Munich *Alarm* referred to is the indicator that the three servo loops (frequency stabilization, interferometer arm difference, interferometer arm length) in the Munich detector are in lock.

It is clear that neither the Glasgow Secondary Feedback nor Secondary Error Point signals, nor the Munich Interferometer Output could be used as vetoes against untrustworthy data, since all of these signals could potentially contain gravitational wave information; and thus any real gravitational waves could veto themselves, which we do not want. Of the remaining data recorded by Watkins, the Multiplexer Synchronisation, Minute Mark and Mains Frequency signals are of no use to our task (though the latter may be useful in a search through the data for continuous wave sources); while the (manually operated) "Lock Regained" signal is thought by the experimenters to be unreliable (Robertson 1992) as the experimenters did not always remember to switch this on at the appropriate moment after regaining lock in the cavity and adjusting any controls that needed to be adjusted.

Further, as can be seen from Table 7.1, there were some signals recorded by the Glasgow experimenters that were not copied (or *retained*) by Watkins's analysis program, which may have turned out to be useful as vetoes at this stage, particularly the Primary Visibility, Low Frequency Feedback, Mains Pulse and Magnetic Pulse signals. This omission is unfortunate, and is a consequence of this being a first attempt at such analysis. Hopefully this will be rectified in future such experimental analyses.

In addition to the above, we have some more possible vetoes that were not recorded by the experimenters, but are based on subsequent analysis. We shall, for convenience, call the

Table 7.1: Data recorded at Glasgow during 100 hour experiment

Signal	Retained	Possible veto
Secondary feedback	$\surd$	
Secondary error point	$\surd$	
Primary error point	$\surd$	$\surd$
Microphone	$\surd$	$\surd$
Secondary visibility	$\surd$	$\surd$
Primary visibility		$\surd$
Seismometer	$\surd$	$\surd$
Low frequency feedback		$\surd$
Oscillation detector		$\surd$
Battery		$\surd$
Multiplexer synchronisation		
Minute mark		
Calibration	$\surd$	$\surd$
Alarm		$\surd$
Mains frequency		
Lock regained		
Mains pulse		$\surd$
Magnetic pulse		$\surd$

Table 7.2: Data recorded at Munich during 100 hour experiment

Signal	Retained	Possible veto
Interferometer output	√	
Table Seismometer	√	√
Suspension seismometer	√	√
Low frequency output	√	√
Alarm	√	√
Laser frequency reference		
Laser frequency correction		
Absolute Armlength photodiode		
Absolute Armlength correction		
Armlength Difference photodiode		
Armlength Difference correction		

above vetoes *primary vetoes*, and those vetoes based on subsequent data analysis *secondary vetoes*. The possible secondary vetoes are as follows:

1. the Glasgow *Gaussian Parameter*,
2. the Munich *Gaussian Parameter*,
3. the wave amplitude discrepancy, and
4. the wave duration discrepancy

(see Section 7.3 for explanations of the above terms).

We shall look first at the primary vetoes.

7.2 Primary vetoes

7.2.1 Method of primary veto selection

As stated above, the function of the vetoes is to remove untrustworthy data and events which are not part of the normal noise behaviour of the detector; without, if possible, removing any real gravitational waves. We could be tempted, therefore, to make a scatter plot of the values of a housekeeping signal against the signal-to-noise output of the detector at those times, and to choose limits on the housekeeping which veto the higher values of both; thus cleaning up the high end of the signal to noise population of both detectors, thus improving the sensitivity of the experiment.

However, this approach is flawed. If we deliberately choose to remove those events of high signal-to-noise, simply because they are inconveniently cluttering the high signal-to-noise end of the distribution, then we are making a biased selection similar to the one RTM made when choosing the optimum threshold for their effect. As with RTM, the result will be to change the distribution of signal-to-noise values in a way which is not understood, and which depends on the whim of the selector, in order to emphasise whatever population the selector wants. Instead, we should use a selection method which is *blind* to the event signal-to-noise, and which depends only on the housekeeping values of the events. Thus, we will have removed those events which cannot be trusted on the basis of the state of detector and laboratory, rather than selecting events which we do not like.

The approach which we shall adopt, then, is to plot a histogram of the values of the housekeeping which occur during the experiment, and to use our knowledge of the housekeeping signals to decide where to draw our line of trust. In particular, where the signal falls into two or more distinct distributions, we shall retain those values which form anything like a Gaussian distribution or which fall close to the mean, while rejecting those which form an outlying group, apparently caused by "something going wrong", or which fall far from the mean.

It would be best if we could return to the original data in each detector and plot a histogram of all of the values taken by each housekeeping stream. However, this task would be extremely lengthy, and would involve, firstly, unpacking all the data again, together with all the Exabyte problems that this would entail, then storing and binning *gigabytes* of real numbers. This procedure is out of the question. The alternative is to look only at those recorded as events by the individual analysis programs. This may well be a biased selection, i.e. the distributions of the values of the housekeeping streams taken at the times of events may not be the same as those taken at times of quiescence in the output of the detector (indeed, it would be very surprising if they were the same[1]), but there is no feasible

[1]In fact, if there were no correlation between housekeeping events and unusual noise events in the detector, there would be no point in using vetoes at all!

alternative. But the procedure of choosing vetoes should still work if the distributions are qualitatively the same, i.e. if there being an event at a certain time does not radically change the probability of a housekeeping stream taking a certain value. Common sense suggests that the distribution of the housekeeping data derived from event or coincident event data will have much the same topology as the distribution for the whole housekeeping stream, but in the former case there will likely be a larger number appearing in the parts of the distribution that we regard as suspicious. If most of the events in the output of the detector are generated by noise which is independent of the housekeeping streams, then this will be the case. We shall proceed as if this last postulate were true, as I believe it is true.[2]

However, even the list of events in each detector is too long to permit binning and histogram plotting routines. Therefore, a random subset of each event list is required. The most convenient random subset of the event lists is the coincidence list, containing, as it does, 83620 entries, roughly evenly spaced through time to fairly represent nonstationarities in the housekeeping streams. This list could not be used if we believed that the event lists were strongly dependent on each other (since then the coincidence list would not be a random subset of the event lists), but such a case would be so far outside our accepted model so far, that we may as well take the independence of the outputs of the two detectors as axiomatic.

7.2.2 Choice of primary vetoes

We shall now examine the histograms of the occurrences of the values taken by the various housekeeping streams. Some housekeeping data come from the Glasgow detector and some from Munich. These subdivide further into those data recorded "point for point", called the *event* data i.e. the housekeeping data recorded at around 1000-10000 Hz, with one value particular to one event in the detector output, and the *group* data, i.e. the variance of a housekeeping stream taken over one group of data. When pertinent, we shall specify whether the data under discussion are of the *event* or *group* type; and we shall not specify either type if we are talking generally about a given housekeeping stream. We shall only go through

[2] In the future, for similar experiments and more advanced data analysis systems, it would be preferable to keep a *running distribution* for each of the housekeeping streams *for all values taken*, say, in the preceding ten minutes, not just for the values at the times of events. Any untrustworthy events could then be vetoed *on line* (i.e. in real time) in a way which is statistically unbiased and which deals naturally with nonstationarities in the distributions of the housekeeping data. Even further in the future, one would hope that there should be no such nonstationarities.

Table 7.3: Values of vetoes adopted from histogram analysis.

Signal	Figure	Unacceptable values
Glasgow Event Primary error pt.	7.1	> 6.0
Glasgow Group Primary error pt.	7.2	> 3250
Glasgow Event Microphone	7.3	no veto
Glasgow Group Microphone	7.4	> 3400
Glasgow Group max Secondary vis.	7.5	< 1900 and > 4100
Glasgow Event Seismometer	7.6	no veto
Glasgow Group Seismometer	7.7	> 2600
Glasgow Calibration	–	applied in analysis
Munich Event Table Seismometer	7.8	no veto
Munich Group Table Seismometer	7.9	> 1500
Munich Event Suspension Seismo.	7.10	no veto
Munich Group Suspension Seismo.	7.11	no veto
Munich Event Low Frequency	7.12	no veto
Munich Group Low Frequency	7.13	> 1500
Munich Alarm	–	applied in analysis

those datastreams with ticks in both columns of Tables 7.1 and 7.2, i.e. those we think are useful *and* have been retained by Watkins's and Nicholson's analysis programs. A list of the selections we shall use for vetoing is given in Table 7.3, followed by the histograms of the behaviour of the housekeeping streams.

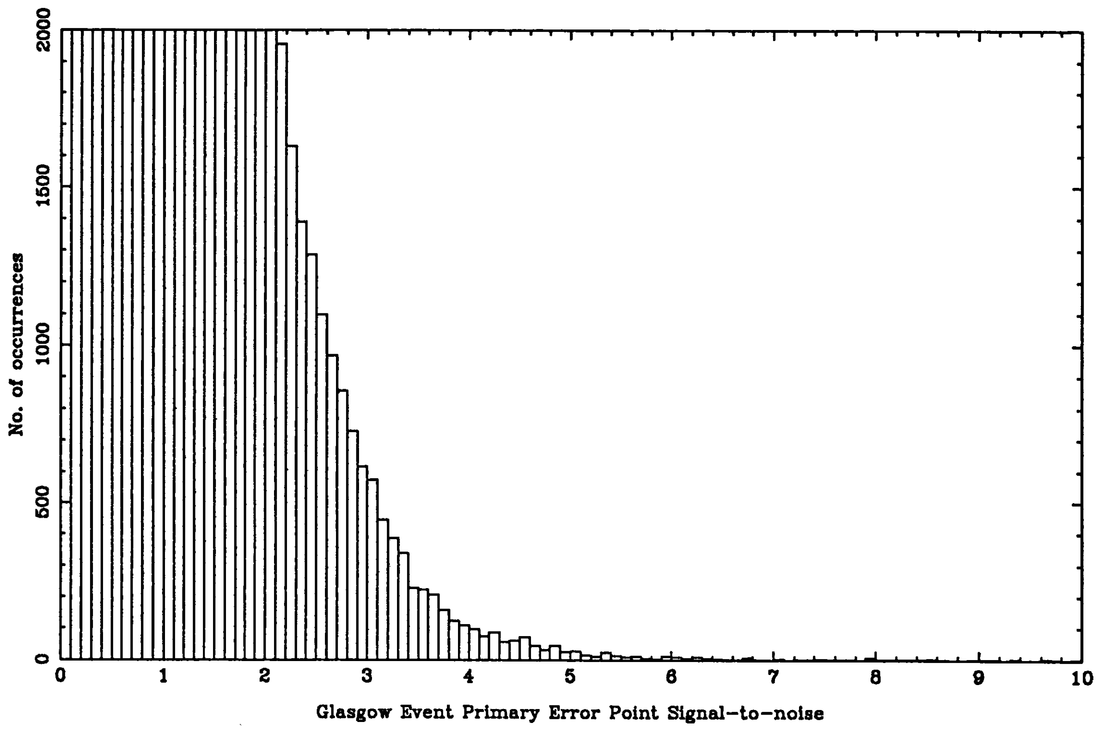


Figure 7.1: Glasgow Event Primary Error Point Signals (coincidence data 5h–93h)

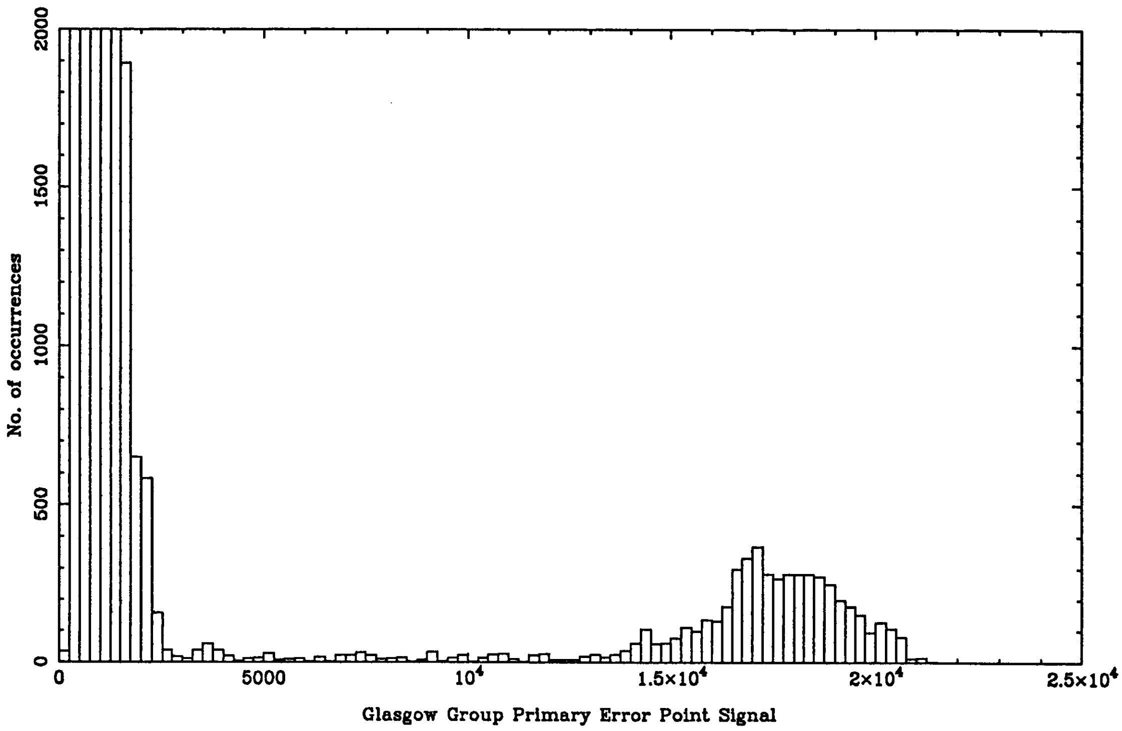


Figure 7.2: Glasgow Group Primary Error Point Signals (coincidence data 5h–93h)

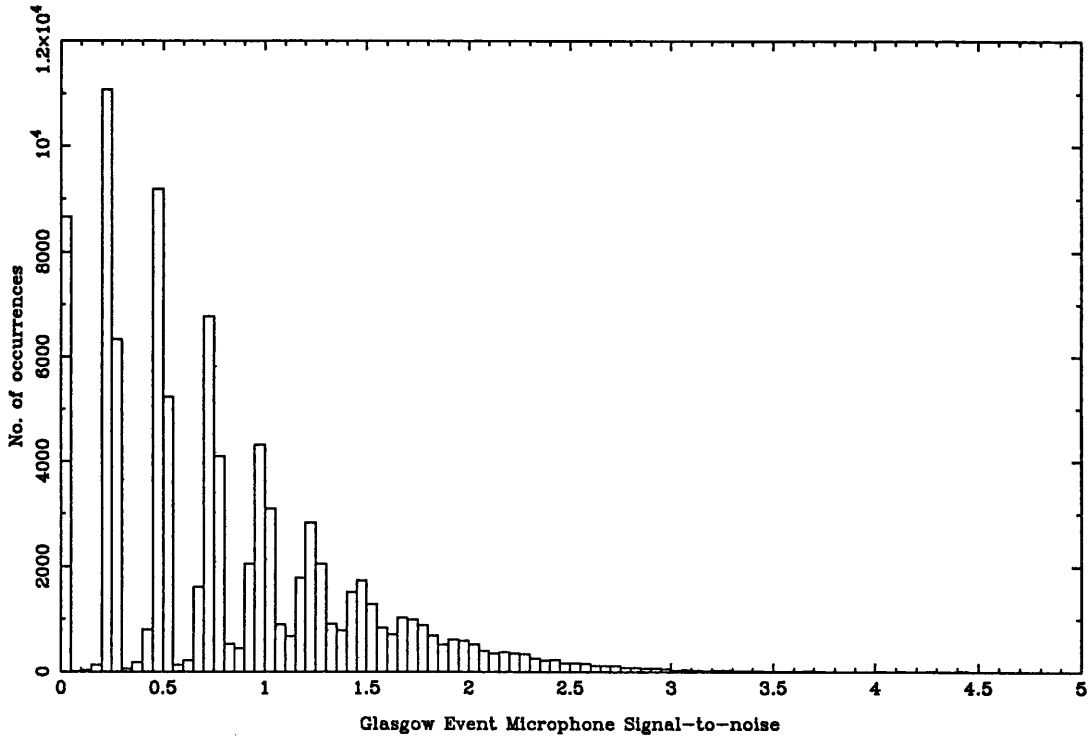


Figure 7.3: Glasgow Event Microphone Signals (coincidence data 5h–93h)

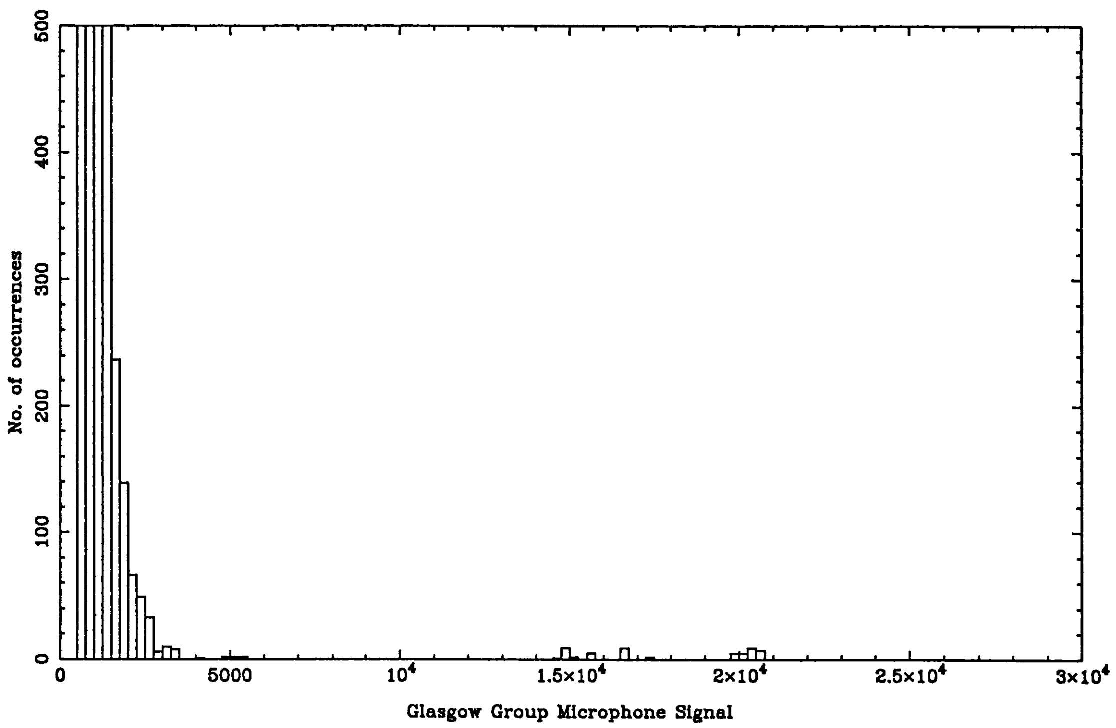


Figure 7.4: Glasgow Group Microphone Signals (coincidence data 5h–93h)

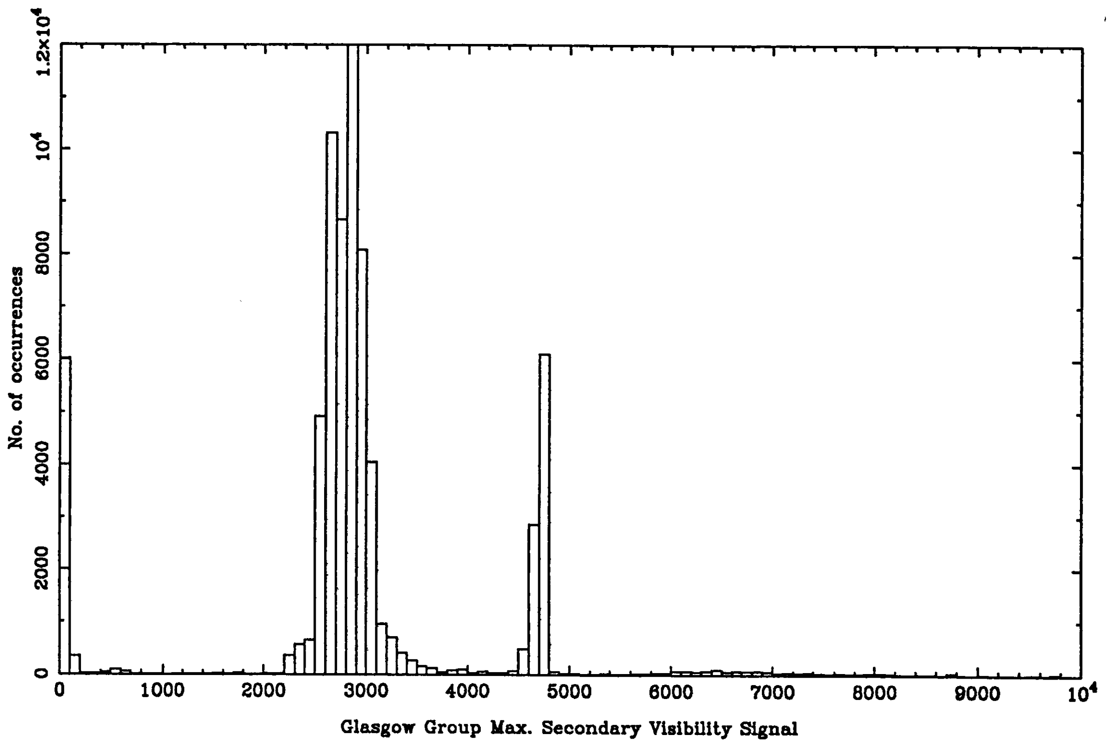


Figure 7.5: Glasgow Group max Secondary Visibility Signals (coincidence data 5h–93h)

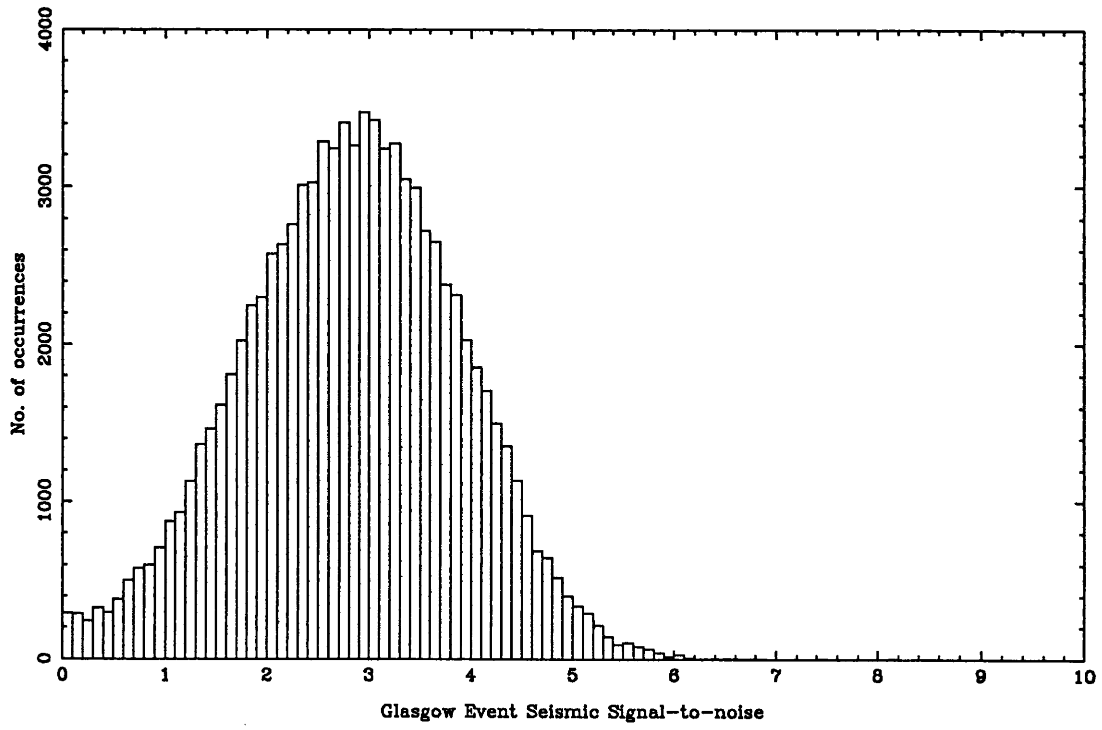


Figure 7.6: Glasgow Event Seismic Signals (coincidence data 5h–93h)

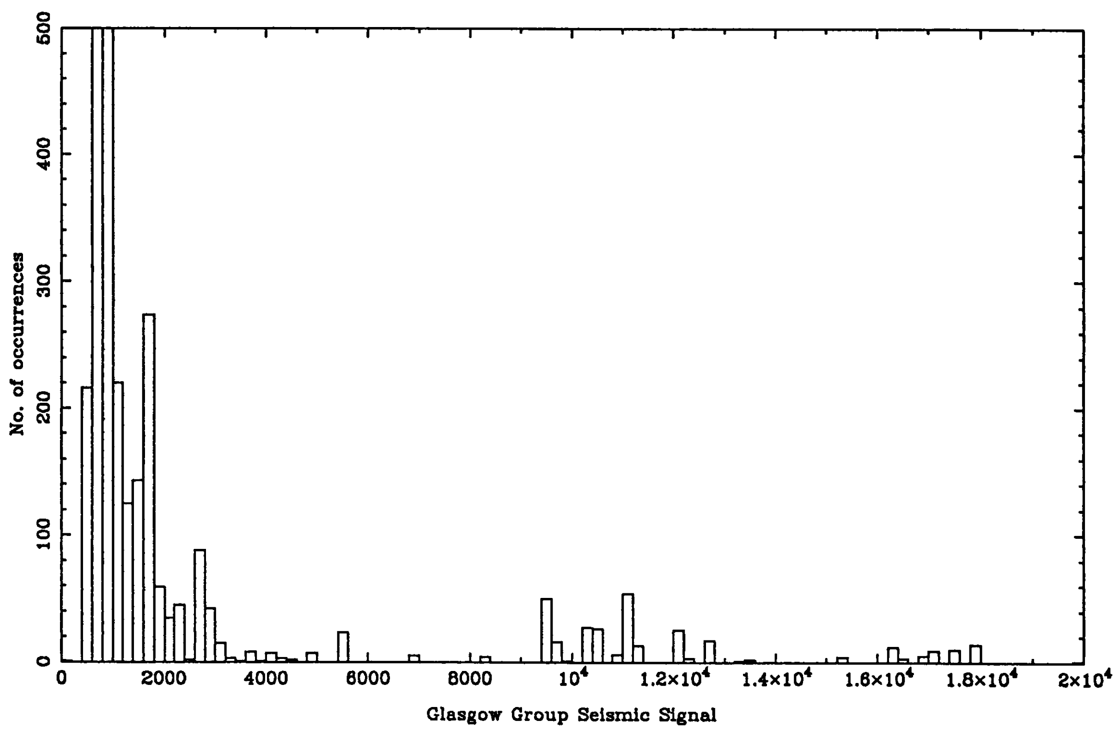


Figure 7.7: Glasgow Group Seismic Signals (coincidence data 5h–93h)

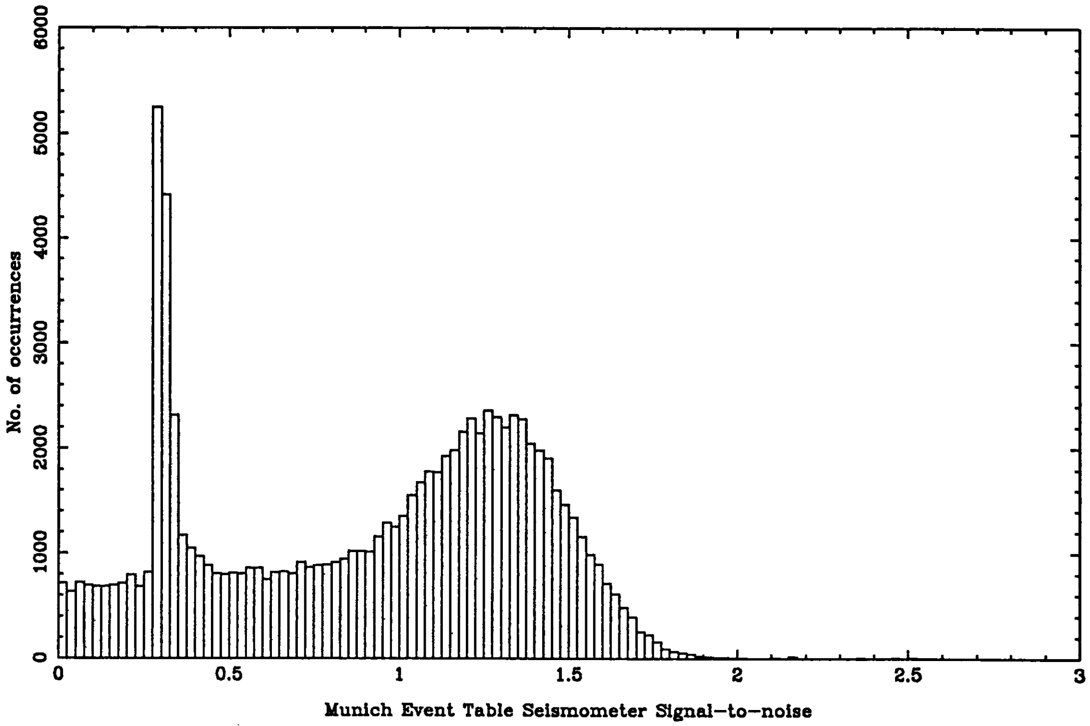


Figure 7.8: Munich Event Table Seismometer Signals (coincidence data 5h–93h)

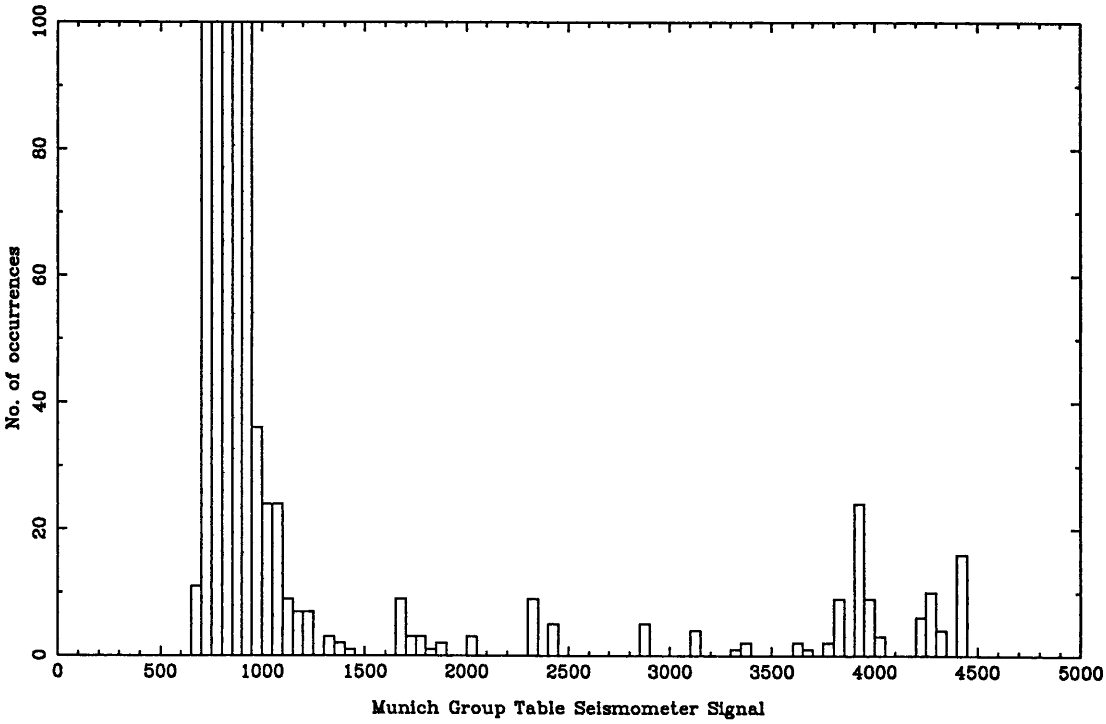


Figure 7.9: Munich Group Table Seismometer Signals (coincidence data 5h–93h)

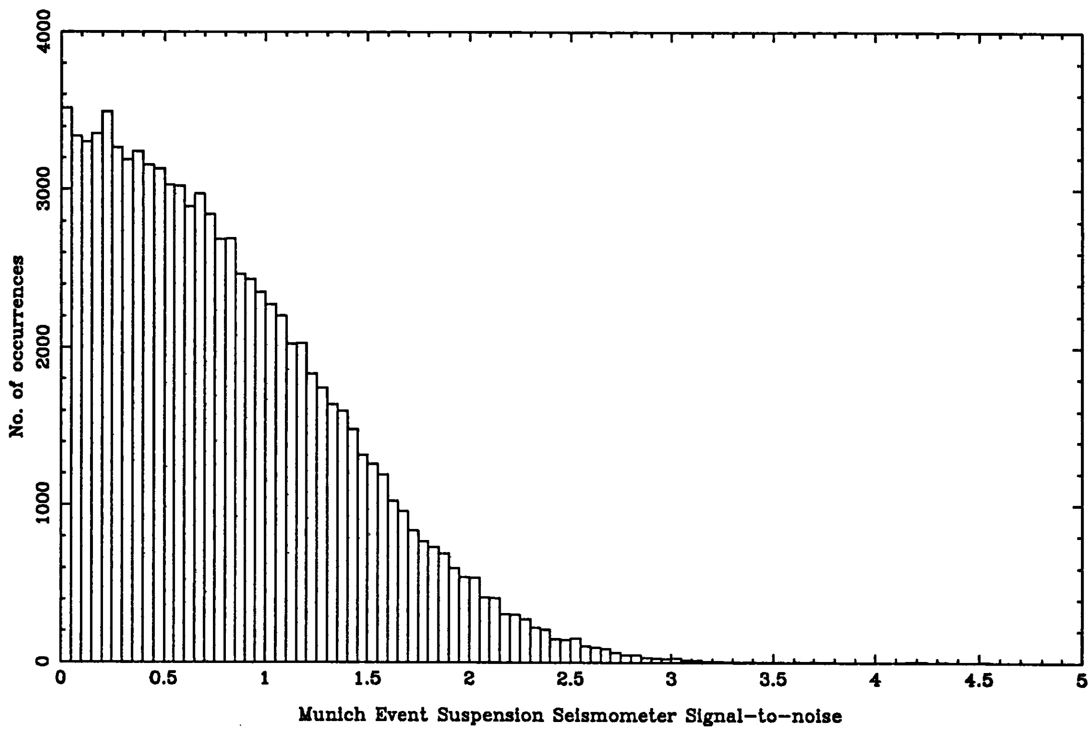


Figure 7.10: Munich Event Suspension Seismometer Signals (coincidence data 5h–93h)

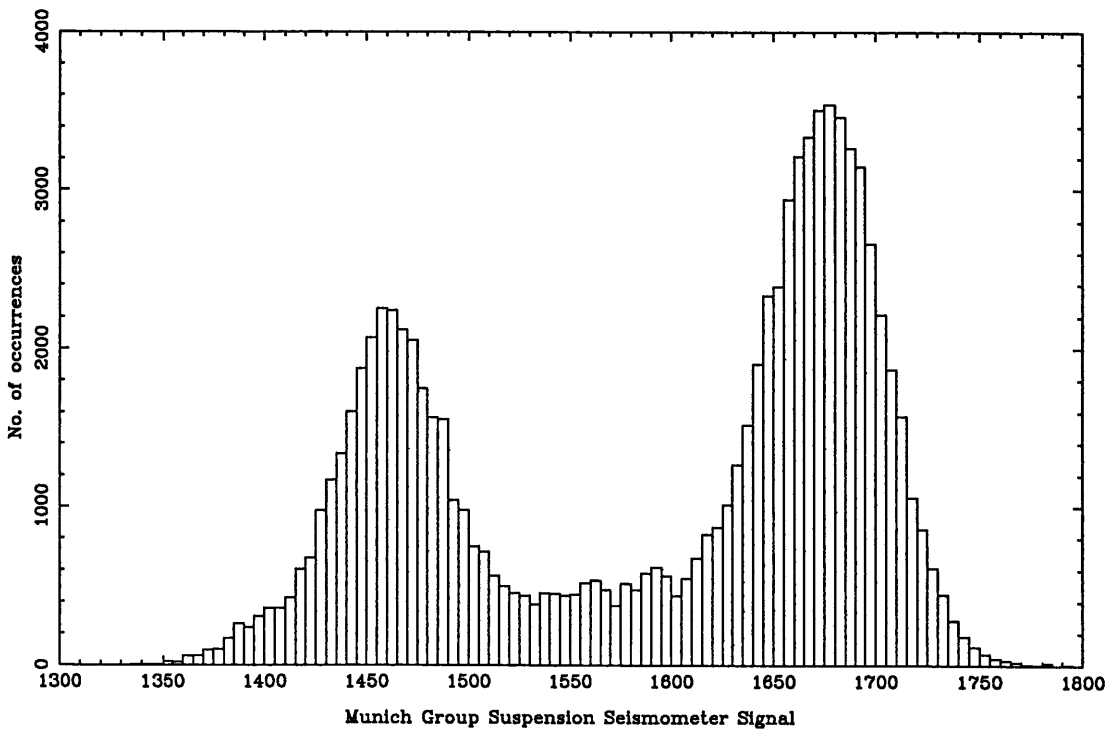


Figure 7.11: Munich Group Suspension Seismometer Signals (coincidence data 5h–93h)

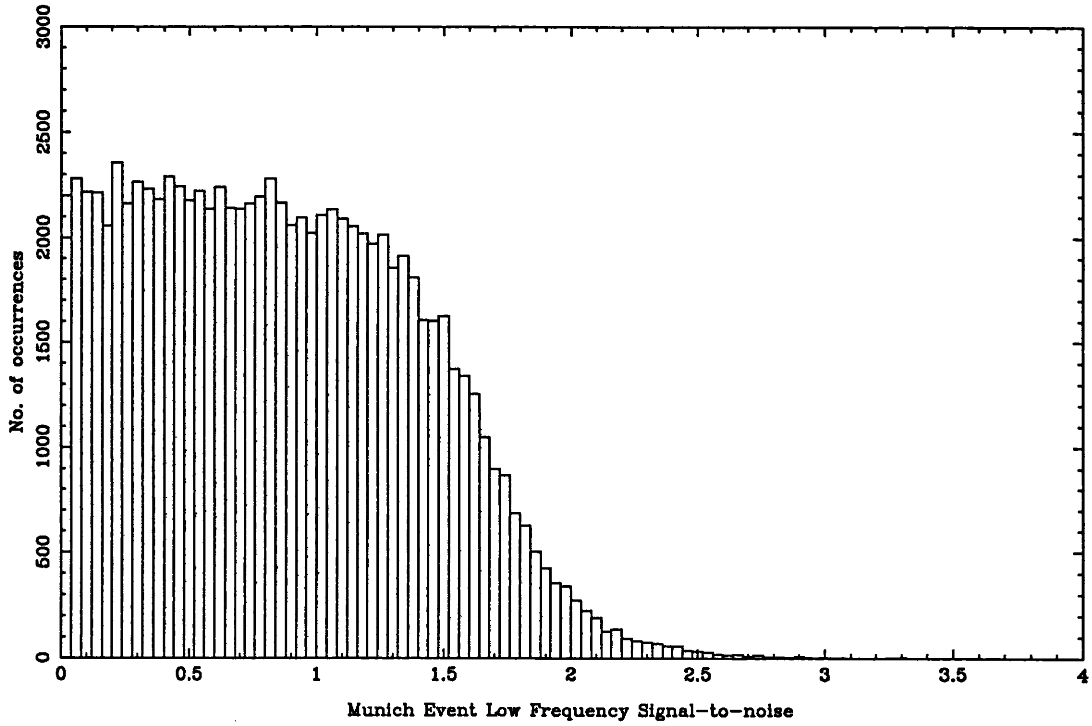


Figure 7.12: Munich Event Low Frequency Signals (coincidence data 5h–93h)

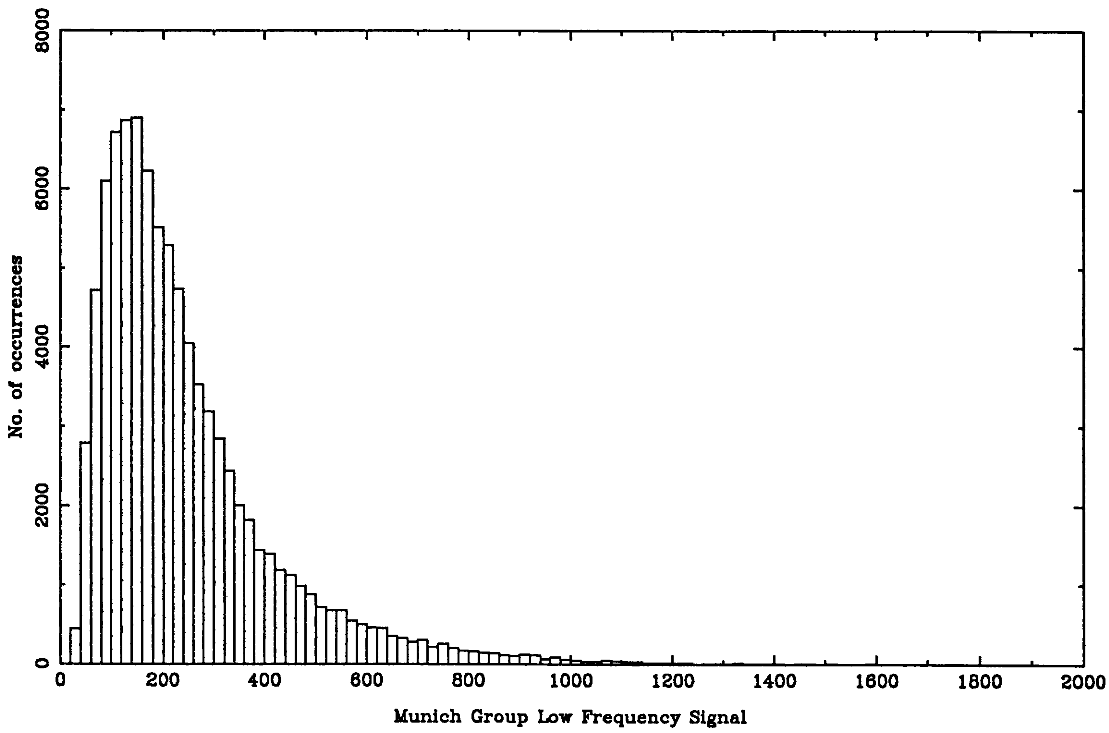


Figure 7.13: Munich Group Low Frequency Signals (coincidence data 5h–93h)

Note:

1. As indicated in Table 7.3, the Glasgow groups containing calibration comb signals and the Munich groups containing Alarm signals had already been vetoed by Watkins's and Nicholson's analysis programs respectively; that is, events occurring during groups containing such signals were not recorded.

2. Any events contained in the first ten datapoints of a Munich group were removed by the coincidence program, for reasons given in Section 6.3.

3. Some of the data streams did not suggest an obvious dividing line for good and bad behaviour, and so I could not set a veto level. These are indicated "no veto" in the right hand column.

4. The Munich Event Table Seismometer signal shows extraordinary behaviour. This I ascribe to the fact that a very strong sawtooth was deliberately added to this signal from time to time in the A/D converter, during the experiment (Rüdiger 1990; Rüdiger 1993).

5. The Munich group suspension seismometer signal appears to have two populations with an overlap area. This could have been used as a veto if one population represented normal behaviour and the other abnormal behaviour. However, as can be seen from Figs. 7.14 and 7.15, the apparently "normal" behaviour of this housekeeping stream seems to vary on a long timescale, indicating nonstationarity of this stream rather than a large number of sporadic seismic perturbations; therefore this stream cannot be used as a reliable veto. The cause of this nonstationarity is unknown. A similar long scale nonstationarity occurs in the Munich detector output noise variance level, σ^2, with a change at around experiment time 15h (not shown here).

6. The other veto levels were chosen in accordance with the method stated in Section 7.2.1.

7.2.3 Initial effects of these choices

The immediate effect of these vetoes is, of course, a reduction in the number of coincidence events tolerated. The numbers of coincidences vetoed independently by each housekeeping stream are given in Table 7.4, at the veto values chosen from Figs. 7.1 to 7.13. There is some

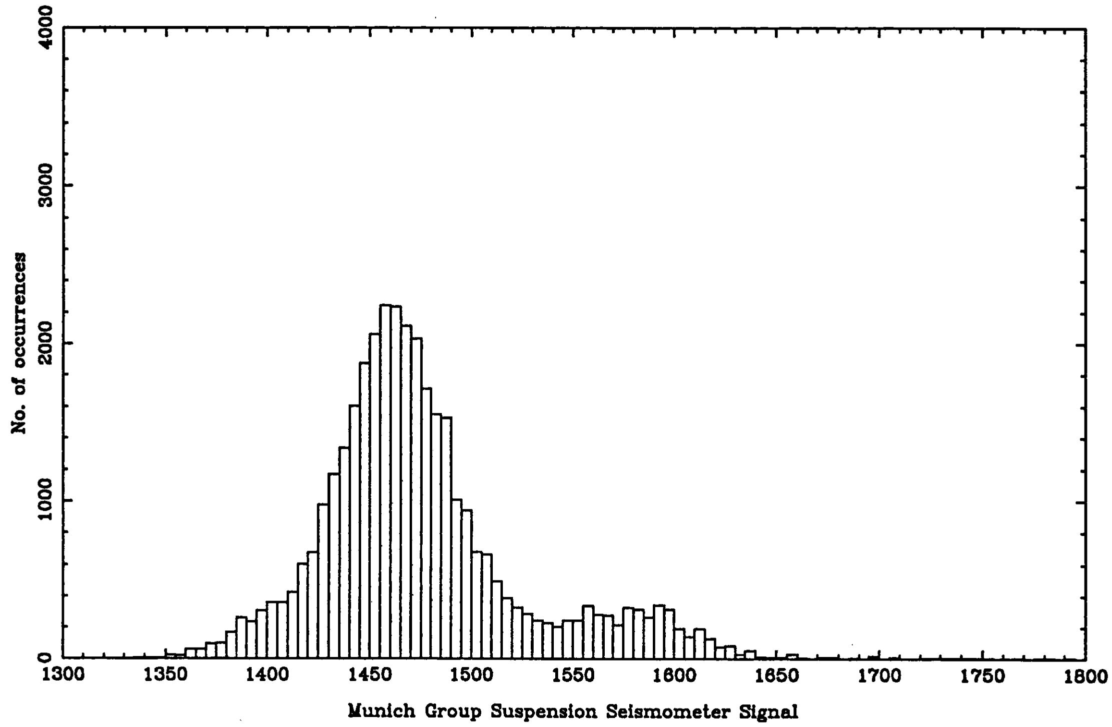


Figure 7.14: Munich Group Suspension Seismometer Signals (coincidence data 5h–46h)

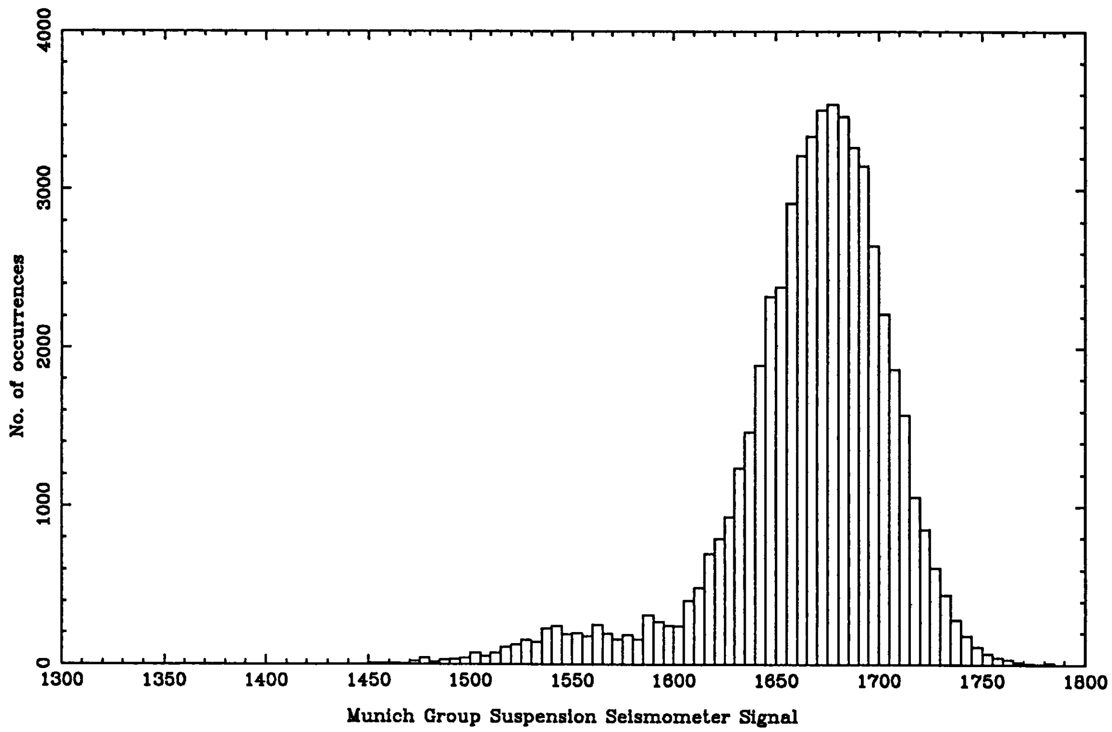


Figure 7.15: Munich Group Suspension Seismometer Signals (coincidence data 46h–93h)

overlap in the numbers of events removed by each of the vetoes, i.e some events are vetoed more than once, so the total number of the events vetoed is somewhat less than the total of the second column of Table 7.4. This is investigated in Sections 7.4.1 and 7.4.2.

Table 7.4: Numbers of coincidence events removed by vetoes

Signal	No. excluded from 83620	Percentage removed
Glasgow Event Primary Error pt.	238	0.3
Glasgow Group Primary Error pt.	5113	6.1
Glasgow Event Microphone	0	0.0
Glasgow Group Microphone	56	0.1
Glasgow Group max Secondary Vis.	17206	20.6
Glasgow Event Seismometer	0	0.0
Glasgow Group Seismometer	511	0.6
Glasgow Calibration	–	applied in analysis
Munich Event Table Seismometer	0	0.0
Munich Group Table Seismometer	123	0.1
Munich Event Suspension Seismo.	0	0
Munich Group Suspension Seismo.	0	0
Munich Event Low Frequency	0	0
Munich Group Low Frequency	12	0.01
Munich Alarm	–	applied in analysis

Prior to the implementation of the vetoes, there were 83620 coincidences recorded during the time of satisfactory detector operation, viz. 5h–93h experiment time. After the primary vetoes were applied, this number went down to 65531, 78% of the original number. This is not, of course, to say that 22% of the *data recorded* have been discounted, only 22% of the *coincidence events* recorded. It is to be expected that, to a certain extent, a larger fraction of the events will be vetoed than fraction of "normal" data. The reason for this is our hypothesis that unusual behaviour in the housekeeping streams will correlate with events in the output of the detector; e.g. a seismic perturbation in the laboratory may be picked up by the seismometer and recorded in the event or group seismic signal, and may also be picked up by the detector, resulting in an event and, in some cases, a coincident event; thus we have a coincident event with a commensurate seismic signal.

For example, as can be seen in Table 7.4, 17206/83620=20.6% of those coincidence events are vetoed on the basis of the Glasgow detector laser cavities being out of lock. Therefore, approximately 20.6% of Glasgow events are vetoed (since the coincident events are a random subset of the Glasgow events; i.e there is no reason to believe that at times of the detector being out of lock, a different fraction of Glasgow events are in coincidence than normal). But this is certainly not the same as saying that 20.6% of Glasgow detector output is vetoed. In fact, the log kept by the Glasgow team tells us that the detector was out of lock approximately 10% of the time. So this tells us that more coincidence events are being created while the detector is out of lock than when it is in lock: this is not surprising since the noise during these periods is much worse than during quiescent periods (see Watkins 1991). Thus around 21% of Glasgow events are created in the 10% of the time. So something like twice as many events are created while the detector is out of lock.

However, the Secondary Visibility is special in that the data recorded while the detector is out of lock are effectively nonsensical; and any real gravitational waves would not even be seen, never mind trusted, due to the drop in sensitivity at these times. The rest of the Glasgow vetoes, including the Gaussian Parameter (see later), are strongly in overlap with the Secondary Visibility signal (see Section 7.4.1); i.e. these are triggered very often when the Secondary Visibility is triggered, and veto little on their own (they veto a total of only 1308 coincidences which the Secondary Visibility does not — see Table 7.6) so only about (1308/83620 $\sim$ 1.6%) of coincident events, hence $\lesssim$1.6% of the data, are vetoed by the Glasgow vetoes while the detector is in lock.

The point I want to make is the following. We have hypothesised that unusual events in the housekeeping correlate with unusual noise in the output of the detector. However, any real gravitational waves which arrive at the detector will not trigger any of the housekeeping signals we have used[3]. Therefore, there will be no correlation between gravitational waves and housekeeping events. This means that while the detectors are working (in lock), if any gravitational waves arrive at the detectors, their chance of accidentally being vetoed by a hamfisted housekeeping stream which is excited at the same time is only about one percent; and such a gravitational wave would never have been trusted anyway.

[3]This is also true, but not obviously so, for the Gaussian Parameter, to which we shall return in 7.3.1.

7.3 Secondary vetoes

7.3.1 The Glasgow Gaussian Parameter

The Gaussian Parameter, Z, was devised by Watkins in his thesis as an indicator of the general state of the detector, and hence the reliability of the data at a particular time. Therein, the Gaussian Parameter in defined as follows: (for more explicit details, the reader is referred to his thesis)

Consider a group of data, 26208 values of $|S/N|$ in the case of Glasgow. We want to know, in some sense, how reliable the data are. One way of doing this is to look at the distribution of the data, and to see how closely it resembles a perfect Gaussian distribution. So bin the values of $|S/N|$ into 500 bins, each of width $\sigma/50$, so that the $|S/N|$ space from 0 to 10σ is divided evenly. Any events of $|S/N|$ exceeding 10σ are put in the last bin. Then count up the number of events falling in the first five bins; call this n_σ. Then the Gaussian Parameter is given by

$$Z = \frac{n_\sigma}{2086}, \tag{7.1}$$

where 2086 is the number expected in the first five bins, were the distribution Gaussian from which the 26208 values were taken (more accurately, $26208 \times \left[2 \times \int_0^{\frac{1}{10}} e^{-x^2/2} dx\right]$).

Clearly, were the underlying distribution of the data Gaussian, we would expect Z to take a value close to 1.0.[4] Large deviations from this value suggest, but do not guarantee, deviations from Gaussian-ness in the underlying distribution. One could be unlucky, for instance, and the sample distribution of the 26208 values could deviate wildly from a Gaussian distribution, while the data were in fact taken from a Gaussian distribution (i.e. the detector was behaving well). However, in general, and with no obvious exceptions, Watkins finds that the distribution of the data is either one or the other; i.e. the data are either approximately Gaussian or are completely wild, with a greatly exaggerated high amplitude tail. He also finds that these periods of "wildness" generally correspond to times when the secondary cavity of the detector is out of lock.

[4]We could also expect the values of Z to have an approximately binomial spread of values about 1.0, since we are counting the number of times out of 26208 that an event of given probability takes place (event: output value falls in first five bins of the whole group), although it is more complicated than this, because the probability of an event occurring depends on σ, which in turn depends on the values of the other 26207 output values.

Watkins does not explicitly investigate confidence in this indicator, but simulations with computer generated data indicate that data taken from a Gaussian distribution will most usually take values of around $Z = 1.00 \pm 0.05$ (see Fig. 7.16). Using 14000 simulated Gaussian datasets of size 26208, we find that none has a Gaussian Parameter of greater than 1.07.

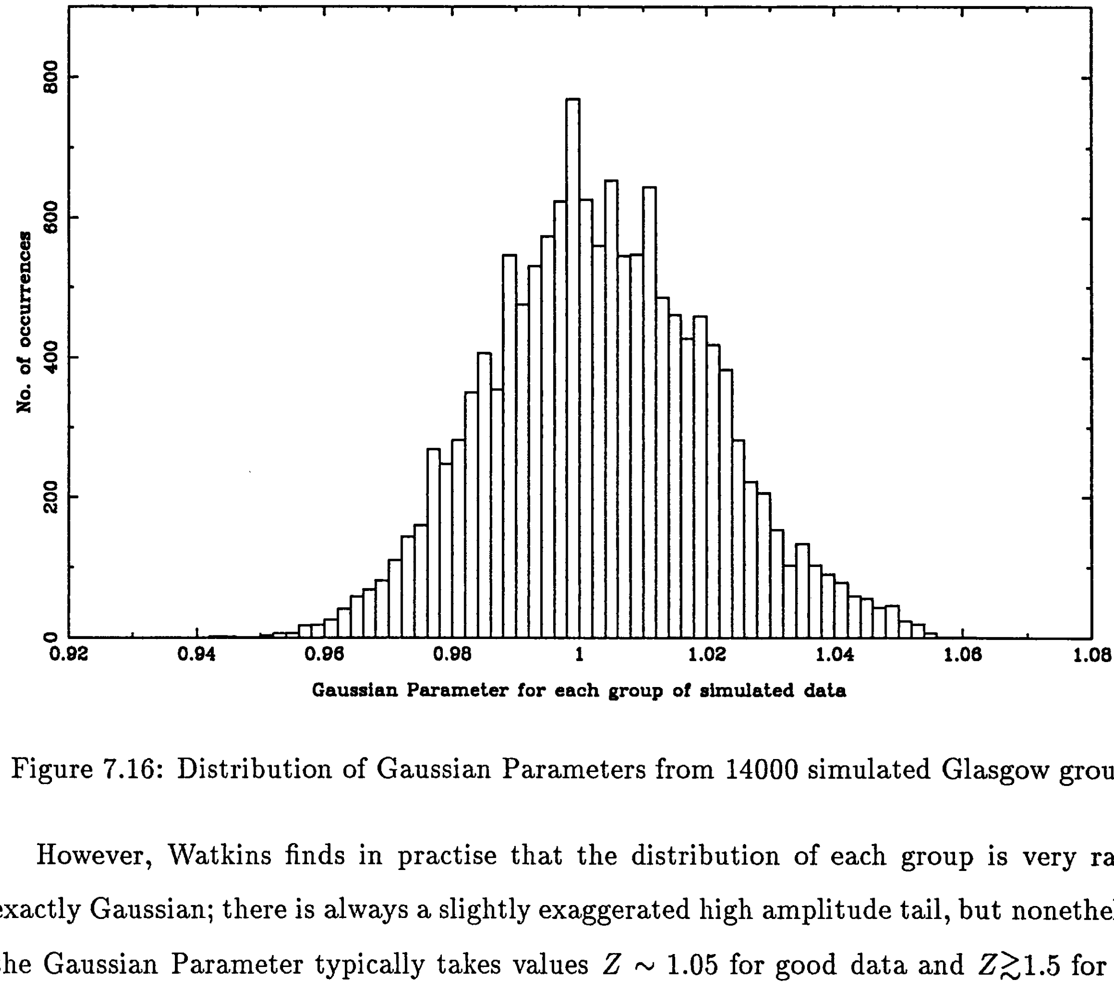


Figure 7.16: Distribution of Gaussian Parameters from 14000 simulated Glasgow groups

However, Watkins finds in practise that the distribution of each group is very rarely exactly Gaussian; there is always a slightly exaggerated high amplitude tail, but nonetheless, the Gaussian Parameter typically takes values $Z \sim 1.05$ for good data and $Z \gtrsim 1.5$ for bad (out of lock) data. Further, this Gaussian Parameter appears to switch between the lower and the higher value as the detector goes in and out of lock (see Fig. 7.17 reproduced from Watkins 1991). Thus, there is reason to believe that the Gaussian Parameter is a good veto for removing badly behaved data, with a cutoff value of around 1.5.

Unfortunately, this thinking is not precise enough for our task. As can be seen from Fig. 7.18, a plot of the Gaussian Parameters taken from the coincidence data, there is no obvious best cutoff for good and bad data.

An immediate corollary of this, which may already have been clear to the reader, is that

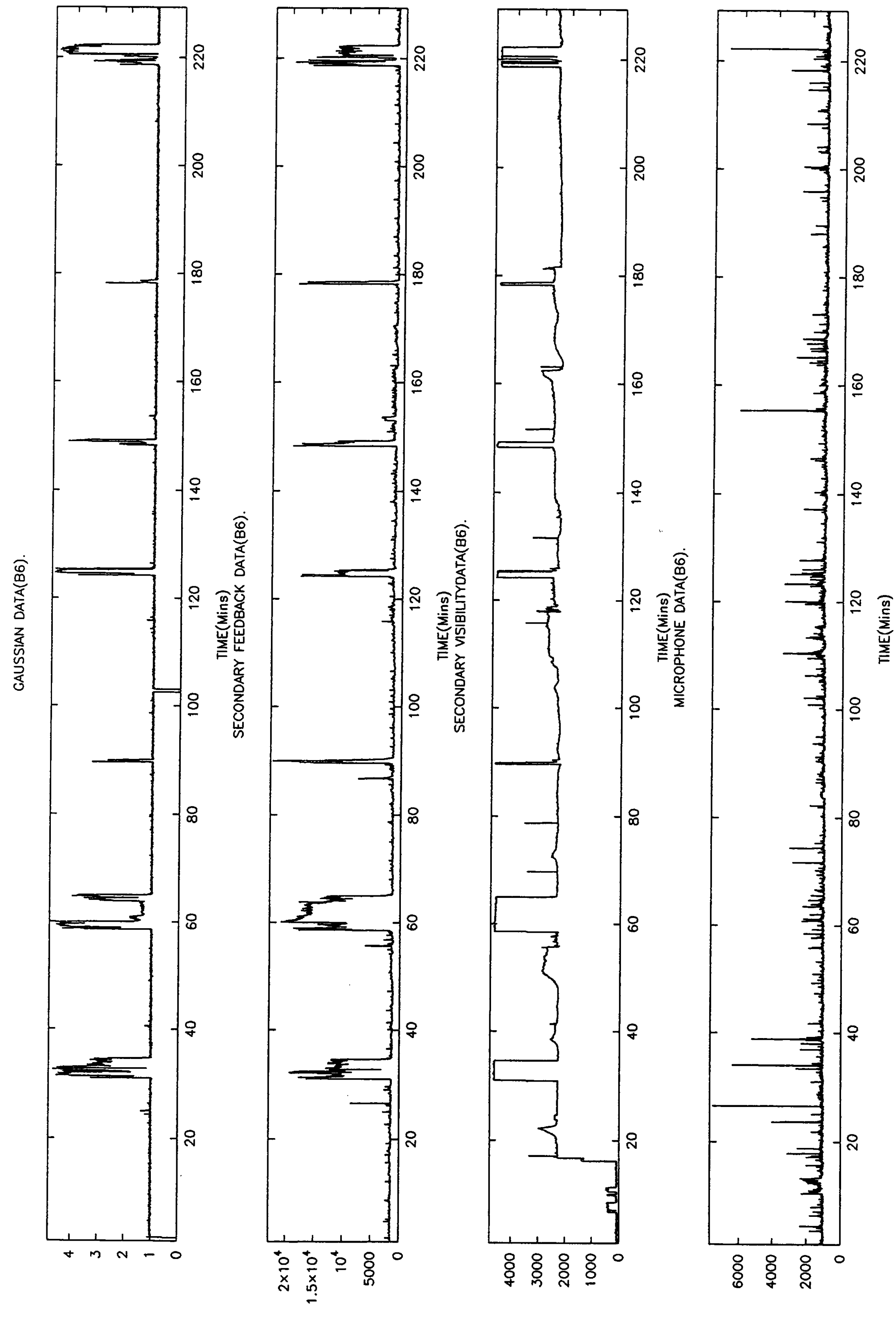


Figure 7.17: Sample Glasgow Output (taken from Watkins 1991)

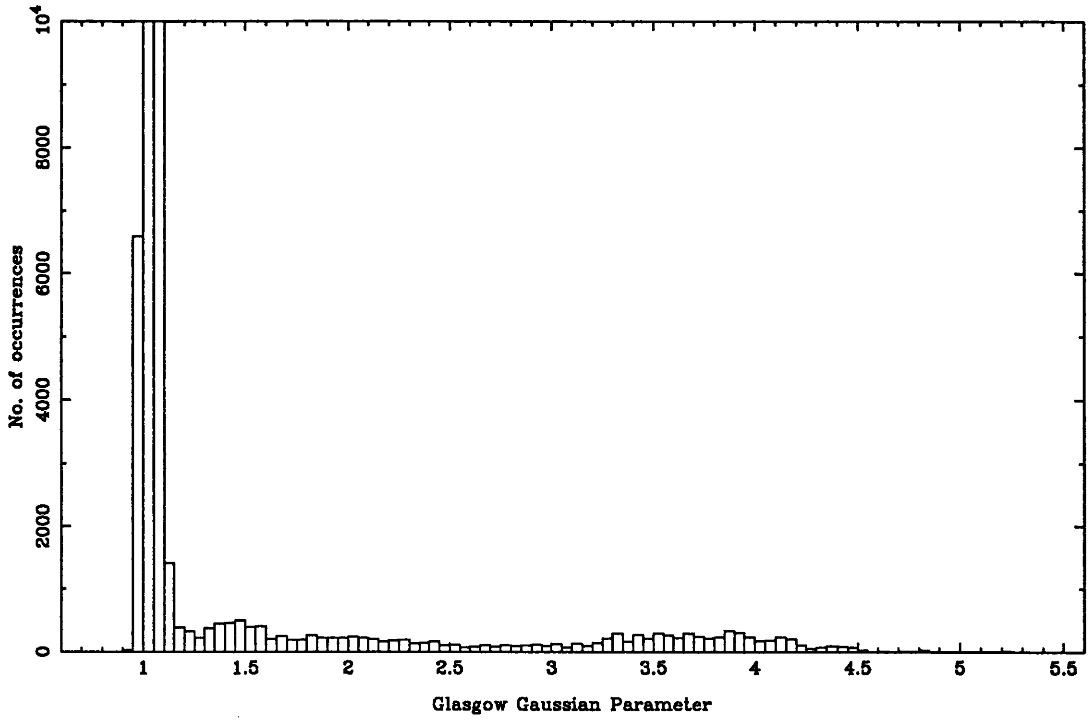


Figure 7.18: Distribution of Glasgow Gaussian Parameters (coincidence data 5h–93h)

for almost any choice of cutoff that we may make, there is a possibility of a group of data taken from a good period (in lock) having, by chance, an unusual sample distribution leading to a high value for Z. It is certainly not obvious that a group coming from a good period but with a randomly-inflicted high Z should be vetoed as "bad data". Further, due to the nature of the Gaussian Parameter, there is the risk of any *real* long, strong gravitational wave events being *self-vetoing*, since the presence of such an event in a group could substantially change the distribution of that group. As an illustration of this effect, see Table 7.5. Here is presented a set of simulated Gaussian noise, 26208 data points, which has an unremarkable Gaussian Parameter of $Z = 1.02$. When one "adds an event" by changing a contiguous subset of datapoints (say, 10 points, corresponding to 0.5ms) to a high S/N value of, say, 10σ, the Gaussian Parameter increases as shown.

Although these artificial events may seem too strong or too long to be physically reasonable, it must be pointed out that the model for expected sources of gravitational waves in our galaxy or in the universe may not be complete. If possible, we should avoid vetoing such very long, very strong unpredicted events which may be there. This is a major disadvantage

Table 7.5: Effect on Gaussian Parameter on introducing an artificial event into the group

Event S/N	Event length (datapoints)	Z
0	0	1.02
10	10	1.036
20	10	1.088
20	20	1.162
20	200	1.97

of the Gaussian Parameter being used as a veto.

The main advantage of the Gaussian Parameter being used as a veto is that it removes events recorded when the distribution of data point amplitudes was very non-Gaussian, but none of the other vetoes is excited outside its acceptable values. In such cases, something is clearly going wrong in the detector, which has not been noted by the other vetoes; and I hypothesise that such data should not be used.

Of course, we don't want to set a veto level so strict that either (a) a large fraction of data not vetoed by other housekeeping streams is so removed, or (b) any possible long, strong gravitational waves would veto themselves. Erring on the side of caution, noting the model for collapse events in our galaxy (and expected amplitudes and durations for said events), and looking at Table 7.5, I have chosen to veto for $Z > 2.0$. With this level set, the Gaussian Parameter for Glasgow removes 7536 of the 83620 coincidences. However, after all other Glasgow vetoes have been applied, only 119 are removed by the Glasgow Gaussian Parameter at this level. These are the events which I hypothesise coincide with the detector malfunctioning, while this has not been picked up by the other vetoes.

Note that, of the 7536 coincidences vetoed by the Gaussian Parameter, nearly all (7103) are also removed by the Secondary Visibility signal. On the other hand, from Table 7.4, the Secondary Visibility veto removes 17206 coincidences, more than twice the number removed by the Gaussian Parameter. So at this level at least, contrary to suggestions I have heard, the Gaussian Parameter is certainly not a replacement for the Secondary Visibility Signal. We shall see shortly, though, that it is probably a fair general indicator of the state of the detector (see Section 7.4.1).

7.3.2 The Munich Gaussian Parameter

In the Munich case, the detector is a delay line interferometer, not a Fabry-Perot cavity like the Glasgow prototype, so there is no need to maintain resonance in the detector arms. However, there can still be a loss of lock in the servos controlling the arm lengths and the laser frequency of the Munich detector. This out-of-lock situation is read by the *Alarm* (Section 7.1) and is vetoed by Nicholson's analysis program before the coincidence analysis stage. Thus, we do not need a Munich Gaussian Parameter to veto these situations. However, we would like the water-cooling events, which did not result in a loss of lock, to be removed. It would be helpful if the Munich Gaussian Parameter fulfilled this role.

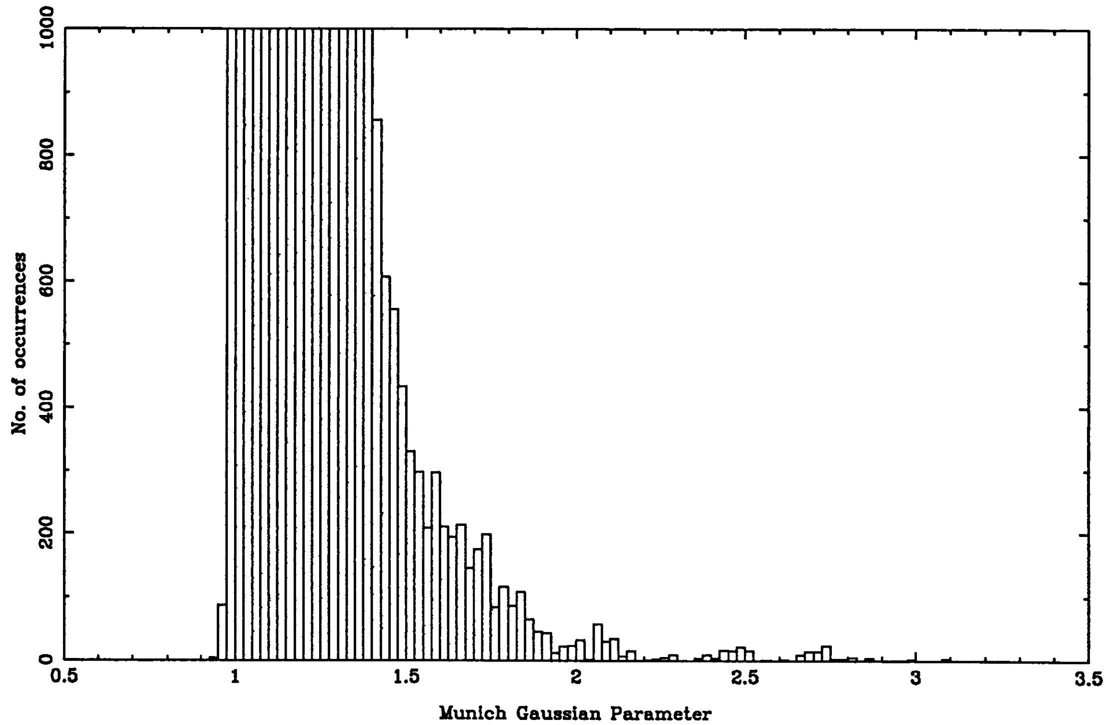


Figure 7.19: Distribution of Munich Gaussian Parameters (coincidence data 5h–93h)

Unfortunately, however, as can be seen from Fig. 7.19, the "water cooling" events do not stand out as a population that can easily be excluded. This is not surprising, though, because their length and amplitude (see Section 6.3) would suggest a Gaussian Parameter of between 1.0 and 1.5 (see Table 7.5). This is confirmed by the time series plot of the Munich Gaussian Parameter for a sample time of 57h 20m to 62h 34m 45s (see later: Fig. 7.22), where one can actually see the regular blips in the Parameter between 2.18×10^5 and 2.24×10^5, spaced about 500 s (8 minutes) apart. These blips only reach a Gaussian Parameter value of about 1.3, and so are too low to be vetoed without the risk of throwing away good data or data containing a real gravitational wave event. They are also too irregular to use their "periodic" (7 minutes $\pm$ 2 minutes) occurrence rate.

The laser refill events, as described elsewhere (Sections 6.3, 7.4.2), would seem to be represented by exceptional double events, like the one at 2.233×10^5 of Fig. 7.22, with Gaussian Parameter values of about 3.6 and 4.3. These events caused a loss of lock (Rüdiger 1990) and will not have been recorded by Nicholson's analysis program. This is mirrored by the fact that the highest value of the Gaussian Parameter seen in the coincidence data is about 3.1 (Figure 7.19). [5]

However, as can be seen in Fig. 7.19, the behaviour of the Munich Gaussian Parameter is a little more helpful than that of Glasgow. In Fig. 7.19, there are several outlying values of the Gaussian Parameter, above $Z \sim 2.0$, which do not belong to the tail of the main distribution. This observation, together with the fact that the length of a Munich data group is close to that of a Glasgow data group (30000 against 26208 datapoints), meaning that Table 7.5 will apply here also; together with the same reasoning as applied to the Glasgow Gaussian Parameter, indicates that vetoing for $Z > 2.0$ is the best procedure in the case of Munich also.

7.3.3 Duration discrepancy

One could argue that any real gravitational waves should have approximately the same duration in time at any detector; therefore one could propose some sort of tolerance criterion which forbids any coincidences which have radically different durations in the two detector outputs. For example, if a coincidence occurred in the two detectors at certain S/N levels, and the events had a duration of 0.1 ms in one detector and 1 ms in another (the *duration* of an event, as recorded in the results of the independent analysis of one of the detector outputs, is the length of time for which the output is above the threshold of 4σ), then this coincidence could be vetoed on the grounds that no real event could have caused this response.

[5] Similarly, none of the other strong peaks in Fig. 7.22 have found their way into the coincidence list suggesting that, whatever their origin, such stong events either (a) have caused loss of lock and have not been recorded by Nicholson's program, or (b) are rare enough that none has coincided with a Glasgow event within the light travel window.

Unfortunately, this is not so. If, as is the case in this experiment, the two detectors have different sensitivities, their thresholds, placed at 4σ, will be at different values of h, the real amplitude. It is h which will be the same at both detectors at all times. The unreliability of this putative veto is demonstrated by the following fairly general example.

Consider a burst event taking place in our galaxy, and suppose its morphology is that shown in Fig. 7.20. Also suppose, for the sake of a simple example, that the noise contributions in the two detectors are small during the event (noise could be added in to give a family of similar examples, each of which is subject to the same argument). Then let the thresholds in the two detectors be as shown in h-space, both being 4σ in signal-to-noise space, but being different in h-space due to the different sensitivities. Here, Glasgow is twice as sensitive as Munich; and so its threshold is at half the level of Munich's, in h-space.

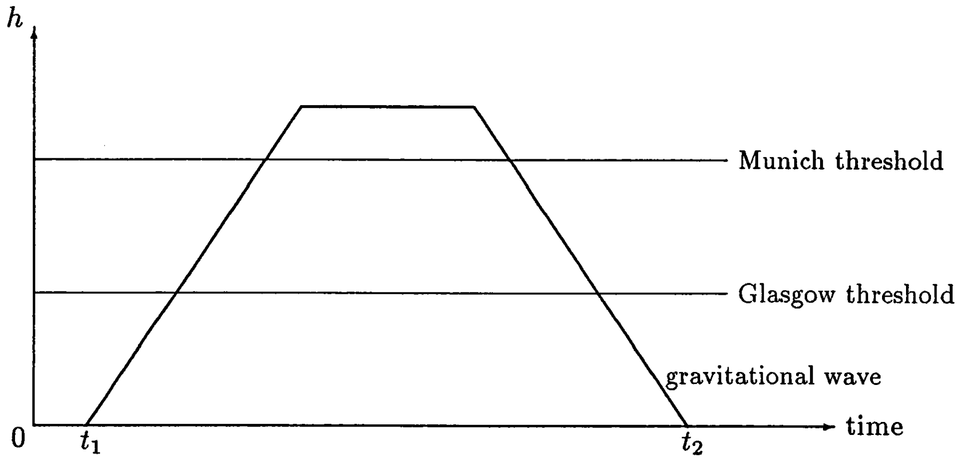


Figure 7.20: Example gravitational wave duration above the threshold of two detectors

With this example event, the signal rises from zero, crosses the two thresholds in turn as its amplitude increases, stays at some plateau level, then decays, and crosses the two thresholds again. As can be seen, the event spends much more time above threshold in the more sensitive detector than in the less sensitive detector. This discrepancy will depend on the sensitivities of the detectors, the orientations of the detectors, the morphology of the event (the longer the ascent and decay, the worse the discrepancy) and the noise contributions from the detectors (which will, of course, be different). Since the sensitivities of the detectors are changing with time, the direction to the source is not known, and the event morphology and noise contributions of the individual detectors are not known a priori, this makes the problem very difficult to invert: given two coincident observations of differing duration, how does one decide whether to tolerate the event? These problems conspire to make any attempt to use such a veto appear extremely unreliable, and so we shall not use this veto.

Note that if, in the future, more information were recorded about the detailed structure of threshold crossing events, we could define the *duration* of an event to be, for example, the duration at half-maximum amplitude. This may then be a useful veto.

7.3.4 Amplitude discrepancy

Theoretically, the discrepancy in amplitude between two coincident observations should be easier to use than the discrepancy in duration, primarily because the amplitude discrepancy of two observations of a real wave is much less dependent on the event morphology, which is not well modelled. The amplitude discrepancy effectively depends only on the sensitivities of and noise contributions in the two detectors (if one ignores the small misalignment of the Glasgow and Munich detectors — see Section 6.2.2).

We shall assume that each of those coincidences crossing threshold is from one of two distributions:

1. bad noise in both detectors ($|S/N| > 4$), and
2. strong gravitational wave event plus noise.

In case 2, we don't know what was the original amplitude of the source wave. However, nearly all such sources would coincide with moderate noise in both detectors ($|S/N| \lesssim 3$) since this is the rule rather than the exception. Further, the contribution of the wave amplitude to the excitation of the detector would be the same in both observations, because the detectors are aligned. Thus we can subtract one observed amplitude from the other, and we are left with a *discrepancy* which is a random variable due only to small independent noise contributions in the two detectors. This has a distribution we can predict theoretically, based on a model of Gaussian noise. In case 1, the discrepancy would not follow such a neat distribution since the observations are caused by independent high amplitude (and probably non-Gaussian) noise in both detectors. This is the signature which will enable us to differentiate between cases 1 and 2, or at least to remove many which are most likely from 1. [6]

[6] As a separate point, if we had some idea of the distribution of expected sources, we could use the observations to infer the most likely original source amplitude, were a strong gravitational wave responsible for the data crossing threshold (a Bayesian calculation). This would be a function of the two observations, weighted by the sensitivities of the detectors and by the expected distribution of source amplitudes.

Now the standard procedures for removing many spurious coincidences would be to use either a false alarm rate (fixing a low probability of accepting a chance spurious coincidence) or a false dismissal rate (fixing a low probability of dismissing a real coincidence); see e.g. Schutz (1991); Davis (1988). Both would involve calculating a *tolerance* for the discrepancy, i.e. the real number t such that

$$\mathrm{P}\{\text{discrepancy} < t\} = a \tag{7.2}$$

in the case of a false alarm probability a, and

$$\mathrm{P}\{\text{discrepancy} > t\} = a' \tag{7.3}$$

in the case of a false dismissal probability a'. Before we even calculate a or a', we would need to specify exactly what we mean by *discrepancy*, which we shall do shortly.

Although many high amplitude coincident noise events are present, most coincidences are around $4 - 5\sigma$ in both detectors. Therefore, if we wanted a sensible false alarm rate, like 1 per 100 hours, we would need to limit our permitted discrepancy to something like $\sim 4\sigma/83620$=a small fraction of one standard deviation. In this case we would also have a very high chance of removing any *real* coincident gravitational wave signals, which would probably by chance differ by of the order of a few sigma. Therefore I shall adopt the false dismissal approach. I shall show at the end of this section that this method is effective at removing many spurious coincidences, although it does not tell us the probability of any of the remaining coincidences arising from pure noise. We shall do the latter task empirically in Chapter 8.

We shall begin by assuming that we are in domain 2 above rather than domain 1, i.e. that any discrepancy in the amplitude is caused by a relatively small noise contribution in each detector. We shall then adopt a Gaussian model for the noise, noting that even though the data streams are somewhat non-Gaussian at high amplitudes, the bulk of the noise contributions causing discrepancies of probability greater than 10^{-3} will satsifactorily be described by a Gaussian model. We shall use our assumption of Gaussian-ness to predict the behaviour of the discrepancy in the measurements. If the actual discrepancy between two observations is not reasonably explained in this model, then we shall assume that the coincidence in fact comes from population 1, and veto it.

Suppose we have two detectors, which record a coincident event. They observe detector outputs s_1 and s_2, measured in dimensionless strain units. Then

$$s_1 = h + n_1, \tag{7.4}$$

and

$$s_2 = h + n_2; \tag{7.5}$$

where h is the amplitude of the source wave, the same as measured by both detectors, and n_1 and n_2 are the Gaussian noise contributions inherent in the detectors at the times of the observations, with zero mean, transformed into strain units. Immediately, we have

$$s_2 - s_1 = n_2 - n_1. \tag{7.6}$$

Let the two detectors measure peak signal-to-noise values of $(S/N)_1$ and $(S/N)_2$ respectively, at times when the dimensionless (frequency-integrated) broadband standard deviations of the noise are σ_1 and σ_2; i.e.,

$$(S/N)_1 = \frac{s_1}{\sigma_1} \tag{7.7}$$

and

$$(S/N)_2 = \frac{s_2}{\sigma_2} \tag{7.8}$$

so that

$$(S/N)_1 = \sigma_1^{-1}(s_1) = \sigma_1^{-1}(h + n_1) = \sigma_1^{-1}(h) + \sigma_1^{-1}(n_1), \tag{7.9}$$

and

$$(S/N)_2 = \sigma_2^{-1}(s_2) = \sigma_2^{-1}(h + n_2) = \sigma_2^{-1}(h) + \sigma_2^{-1}(n_2). \tag{7.10}$$

Now we have the problem that only the absolute values of S/N were recorded in the individual event analysis (see Section 5.2.1) and hence in the coincidence lists. To proceed, we must assume that $(S/N)_1$ and $(S/N)_2$ have the same sign. Certainly, we will tolerate coincidences which we should veto, but we don't have the sign information to hand anyway. In any case, such spurious coincidences could still be vetoed later by checking them individually. Also, at least this assumption will not cause us to reject any real coincidences, which is the thrust of our false dismissal analysis. Then the following holds,

$$|a|(S/N)_2| - b|(S/N)_1|| = |a(S/N)_2 - b(S/N)_1|, \tag{7.11}$$

for all $a, b > 0$. In particular, we have,

$$\left| |(S/N)_2| - \left(\frac{\sigma_1}{\sigma_2}\right) |(S/N)_1| \right| = \left| (S/N)_2 - \left(\frac{\sigma_1}{\sigma_2}\right) (S/N)_1 \right| \tag{7.12}$$

$$= \left| \sigma_2^{-1}(h) + \sigma_2^{-1}(n_2) - \left(\frac{\sigma_1}{\sigma_2}\right) \sigma_1^{-1}(h) - \left(\frac{\sigma_1}{\sigma_2}\right) \sigma_1^{-1}(n_1) \right| \tag{7.13}$$

by Eq. 7.9 and 7.10,

$$= \left| \sigma_2^{-1}(n_2) - \left(\frac{\sigma_1}{\sigma_2}\right) \sigma_1^{-1}(n_1) \right|. \tag{7.14}$$

Now, as we supposed, the noise contributions, n_1 and n_2 in h-space, have Gaussian distributions; so that we define two standard normal random variables u and v such that

$$u = n_1/\sigma_1 = \sigma_1^{-1}(n_1) \tag{7.15}$$

and

$$v = n_2/\sigma_2 = \sigma_2^{-1}(n_2); \tag{7.16}$$

so that Eq. 7.14 becomes

$$\left| |(S/N)_2| - \frac{\sigma_1}{\sigma_2} |(S/N)_1| \right| = \left| v - \frac{\sigma_1}{\sigma_2} u \right|. \tag{7.17}$$

This equation looks almost tautological, and in fact it almost is in the case of there being no gravitational waves present, since then u and v are $(S/N)_1$ and $(S/N)_2$ by definition. But the equation also tells us what happens when there is a gravitational wave present; and since we have subtracted off any gravitational waves which may be present, we have a relationship between only the observations on the left hand side and a prediction of the noise distribution on the right hand side. We call the quantity on the L.H.S. the *normalised discrepancy* of the two observations, since the term σ_1/σ_2 serves to normalise the output of detector 1 to that of detector 2. If a measured normalised discrepancy on the L.H.S. is unlikely to have arisen during well-behaved Gaussian noise such as the distribution predicted on the R.H.S., then we are probably not in domain 2, and we shall veto it.

Now we want to remove as many spurious coincidences as possible, but to have as low a chance as possible of removing any real, strong coincidences. As a compromise, we shall reject any coincidences which have a normalised discrepancy with probability of less than 10^{-3} of arising during well behaved Gaussian noise, whether containing a real signal or not. This choice was made a priori. It is to be hoped that this will be sufficient to remove many coincidences which to the eye appear clearly spurious and we shall see later that it is so (see Chapter 8). Thus we veto a coincidence if

$$\left| |(S/N)_2| - \frac{\sigma_1}{\sigma_2} |(S/N)_1| \right| > t \tag{7.18}$$

where t is the *tolerance* of the normalised discrepancy, whose value we obtain from

$$P\left\{ \left| |(S/N)_2| - \frac{\sigma_1}{\sigma_2} |(S/N)_1| \right| > t \right\} = 10^{-3} \tag{7.19}$$

where P indicates probability of the event in brackets, based on our model of Gaussian noise; i.e., from 7.17,

$$P\left\{ \left| v - \frac{\sigma_1}{\sigma_2} u \right| > t \right\} = 10^{-3}, \tag{7.20}$$

i.e.,

$$P\left\{ \left[v - \frac{\sigma_1}{\sigma_2} u \right] > t \ \cup \ \left[v - \frac{\sigma_1}{\sigma_2} u \right] < -t \right\} = 10^{-3}. \tag{7.21}$$

To solve Eq. 7.21 for t, we must integrate the joint probability density function, $p(u, v)$, over the shaded region of the u-v diagram Fig. 7.21.

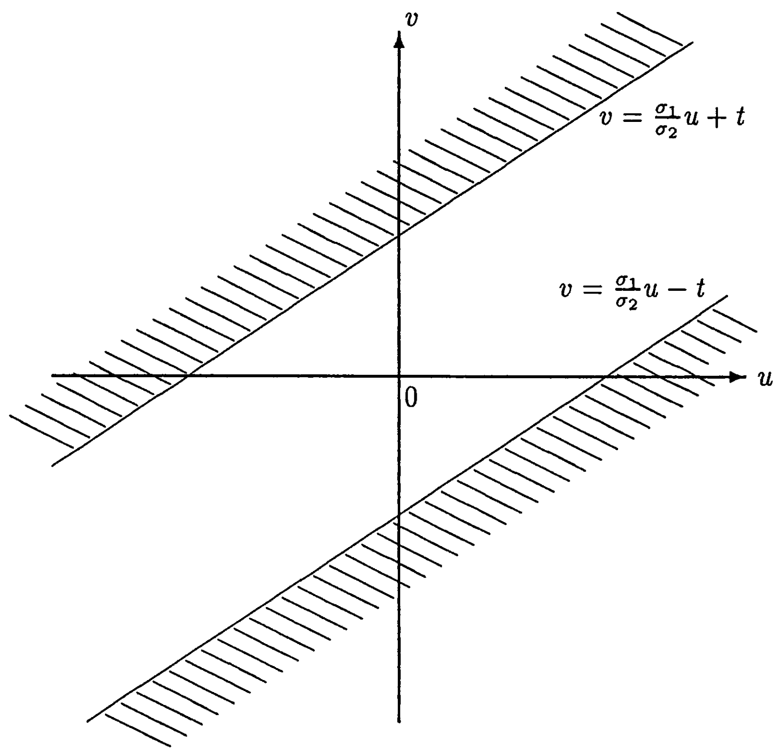


Figure 7.21: Area of integration over which we shall veto a coincidence with discrepant gravitational-wave amplitude

Now by independence of the noise contributions u and v, the joint probability density function $p(u,v)$ is given simply in terms of the individual probability density functions by

$$\begin{aligned} p(u,v) &= p(u)p(v) && (7.22) \\ &= \frac{1}{\sqrt{2\pi}}e^{-u^2/2}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}, && (7.23) \end{aligned}$$

by hypothesis. Hence,

$$\begin{aligned} P\left\{\left|v-\frac{\sigma_1}{\sigma_2}u\right|>t\right\} &= \int_{u=-\infty}^{\infty}\int_{v=\frac{\sigma_1}{\sigma_2}u+t}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\,dvdu \\ &\quad+\int_{u=-\infty}^{\infty}\int_{v=-\infty}^{\frac{\sigma_1}{\sigma_2}u-t}\frac{1}{\sqrt{2\pi}}e^{-u^2/2}\frac{1}{\sqrt{2\pi}}e^{-v^2/2}\,dvdu && (7.24) \\ &= I_1+I_2 \text{ (say)} && (7.25) \\ &\equiv 10^{-3}. && (7.26) \end{aligned}$$

This equation can be solved numerically for t in terms of σ_1 and σ_2, but we shall see that it is more instructive and elegant to proceed analytically. We shall deal firstly with I_1.

We now make our first change of variable, i.e., put

$$\begin{aligned} u &= r\cos\theta \\ v &= r\sin\theta && (7.27) \end{aligned}$$

where r and θ are the usual polar coordinates centred on the origin. We wish to integrate with respect to r first, so the area of the plane in which we are interested is covered by a set of lines radiating from the centre, each of constant θ, beginning at the lines delimiting the area of integration and going off to infinity. If we integrate with respect to r first, then we want the limits of r as a function of θ. Substituting equations 7.27 in the equation of the line demarking the upper region, viz. $v=\frac{\sigma_1}{\sigma_2}u+t$, gives the lower limit of r, viz. $r_{min} = t/(\sin\theta-(\sigma_1/\sigma_2)\cos\theta)$. The delimiting values of θ will be those parallel to the line delimiting the region, which from simple geometry are $\theta_{min}=\tan^{-1}(\frac{\sigma_1}{\sigma_2})$ and $\theta_{max}=\tan^{-1}(\frac{\sigma_1}{\sigma_2})+\pi$. The value of the Jacobian is r. Then,

$$\begin{aligned} I_1 &= \frac{1}{2\pi}\int_{\theta=tan^{-1}(\frac{\sigma_1}{\sigma_2})}^{tan^{-1}(\frac{\sigma_1}{\sigma_2})+\pi}\int_{r=\frac{t}{\sin\theta-\frac{\sigma_1}{\sigma_2}\cos\theta}}^{\infty} e^{-r^2/2}\,rdrd\theta && (7.28) \\ &= \frac{1}{2\pi}\int_{\theta=tan^{-1}(\frac{\sigma_1}{\sigma_2})}^{tan^{-1}(\frac{\sigma_1}{\sigma_2})+\pi} e^{-\frac{t^2}{2(\sin\theta-\frac{\sigma_1}{\sigma_2}\cos\theta)^2}}\,d\theta && (7.29) \end{aligned}$$

$$= \frac{1}{2\pi} \int_{\theta=tan^{-1}(\frac{\sigma_1}{\sigma_2})}^{tan^{-1}(\frac{\sigma_1}{\sigma_2})+\pi} e^{-\frac{\sigma_2^2 t^2}{2(\sigma_2 \sin\theta - \sigma_1 \cos\theta)^2}}\, d\theta \tag{7.30}$$

Looking at the above, especially the limits of integration, it seems natural to look for a substitution in the form of a rotation of angle $\tan^{-1}\sigma_1/\sigma_2$ about the origin. The following was suggested by B. Schutz: the term $\sigma_2 \sin\theta - \sigma_1 \cos\theta$ is of the form $\alpha \sin\theta \cos\phi - \alpha \cos\theta \sin\phi$, with $\alpha \cos\phi = \sigma_2$ and $\alpha \sin\phi = \sigma_1$. Then $\phi = \tan^{-1}(\sigma_1/\sigma_2)$, $\alpha = \sqrt{\sigma_1^2 + \sigma_2^2}$. But $\alpha \sin\theta \cos\phi - \alpha \cos\theta \sin\phi = \alpha \sin(\theta - \phi)$. So we obtain

$$I_1 = \frac{1}{2\pi} \int_{\theta=\phi}^{\theta=\phi+\pi} e^{-\frac{\sigma_2^2 t^2}{2\alpha^2 \sin^2(\theta-\phi)}}\, d\theta. \tag{7.31}$$

Now put $\psi = \theta - \phi$; $d\psi = d\phi$. Then

$$I_1 = \frac{1}{2\pi} \int_0^{\pi} e^{-\frac{\sigma_2^2 t^2}{2(\sigma_1^2+\sigma_2^2)\sin^2\psi}}\, d\psi. \tag{7.32}$$

So I_1 is a function of only one simple unknown parameter, d, say, given by

$$d^2 = \frac{\sigma_2^2 t^2}{2(\sigma_1^2 + \sigma_2^2)} = \frac{t^2}{2\left(1 + \frac{\sigma_1^2}{\sigma_2^2}\right)}; \tag{7.33}$$

and so

$$I_1 = \frac{1}{2\pi} \int_0^{\pi} e^{-\frac{d^2}{\sin^2\psi}}\, d\psi. \tag{7.34}$$

Similarly,

$$I_2 = \frac{1}{2\pi} \int_{\pi}^{2\pi} e^{-\frac{d^2}{\sin^2\psi}}\, d\psi. \tag{7.35}$$

Now $e^{-\frac{d^2}{\sin^2\psi}} = \left[e^{-\frac{1}{\sin^2\psi}}\right]^{d^2}$, and so the integrand of 7.34 repeats for each increment of π. Thus, happily,

$$I_1 = I_2. \tag{7.36}$$

In addition, the integrand is symmetrical about $\psi = \pi/2$. So Eq. 7.34 becomes

$$I_1 = \frac{1}{\pi} \int_0^{\frac{\pi}{2}} e^{-\frac{d^2}{\sin^2\psi}}\, d\psi. \tag{7.37}$$

Now we make a last change of variable. This was suggested by A. Lobo. Put

$$p = \cot\psi \; ; \; dp = -\frac{1}{\sin^2\psi} d\psi. \tag{7.38}$$

Then $\sin^2\psi = 1 + p^2$; and so Eq. 7.37 becomes

$$I_1 = -\frac{1}{\pi} \int_{\infty}^{0} \frac{e^{-d^2(1+p^2)}}{1+p^2} dp \tag{7.39}$$

$$= \frac{e^{-d^2}}{\pi} \int_0^{\infty} \frac{e^{-d^2 p^2}}{1+p^2} dp. \tag{7.40}$$

We can re-express this in terms of the error function, erf. We define

$$\text{erf}\, d = \frac{2}{\sqrt{\pi}} \int_0^d e^{-q^2} dq; \tag{7.41}$$

and the following standard result can be used (see Abramowitz & Stegun 1972):

$$\int_0^\infty \frac{e^{-ap^2}}{p^2 + x^2} dp = \frac{\pi}{2x} e^{ax^2} \text{erfc}\, \sqrt{a}x; \quad (a, x > 0) \tag{7.42}$$

where erfc is the complimentary error function i.e. erfc $\sqrt{a} = 1 - \text{erf}\, \sqrt{a}$. Put $a = d^2$, $x = x^2 = 1$. Then Eq. 7.40 becomes

$$I_1 = \frac{e^{-d^2}}{\pi} \left[\frac{\pi}{2} e^{d^2} \text{erfc}\, d \right] \tag{7.43}$$

$$= \frac{1}{2}[1 - \text{erf}\, d]. \tag{7.44}$$

Then Eq. 7.26, 7.36, and 7.44 give

$$1 - \text{erf}\, d = 10^{-3}. \tag{7.45}$$

Standard tables give $d = 2.323$ i.e.

$$\frac{t^2}{2\left(1 + \frac{\sigma_1^2}{\sigma_2^2}\right)} = d^2 = 5.40, \tag{7.46}$$

by 7.33; i.e.

$$t = 3.29 \left(1 + \frac{\sigma_1^2}{\sigma_2^2}\right)^{\frac{1}{2}}; \tag{7.47}$$

where t is the tolerance defined earlier. So by Eq. 7.18, we reject a coincidence if

$$\left| |(S/N)_2| - \frac{\sigma_1}{\sigma_2} |(S/N)_1| \right| > 3.29 \left(1 + \frac{\sigma_1^2}{\sigma_2^2}\right)^{\frac{1}{2}}. \tag{7.48}$$

We can express this in a way which demonstrates the symmetry of the problem, i.e.,

$$\boxed{||(S/N)_2|\sigma_2 - |(S/N)_1|\sigma_1| > 3.29(\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}} \tag{7.49}$$

Heuristically, this says that of those coincidences thrown away by the h veto, their measured amplitudes are so discrepant that the probability of their discrepancy having been caused by well behaved Gaussian noise of zero mean and variances σ_1^2 and σ_2^2 is less than 10^{-3}. Since gravitational waves are rare, we would expect that any strong signals present would coincide with good noise, being much more common than bad noise, and would not satisfy the inequality. Hence any coincidences satisfying the inequality are very likely spurious and should be rejected.

Clearly, if one would prefer a false dismissal probability other than 10^{-3}, a simple change could be made to Eq. 7.45 and hence to Eq. 7.49.

At this point, a word of warning must be given. All of the preceding vetoes, including the Gaussian Parameter vetoes, were applied independently to the two data streams, and so provided the housekeeping streams of Glasgow and Garching were independent (events in the "seismic" signal, for example, are caused by local vibrations of the laboratory and apparatus, not tectonic motions as the name might suggest), the two data streams surviving the vetoes will be independent. However, this is not so for the h veto described above. The coincidence list surviving the h veto will be a list of pairs of events, which have been selected out in a way where the two events of each pair are in some way dependent on each other. Strictly, therefore, one should not make a calculation of the probability of these coincidences, which calculation is based on the individual S/N distributions of the two detector surviving after the h veto was applied. Nonetheless, the calculation of their probability based on the noise distributions before the h veto was applied will be correct, and this will give an upper limit on the probability on any coincidences which also survive the h veto. Crudely, we could multiply this upper limit probability by the fraction of coincidences which survive the h veto to get an estimate which is closer to the real probability of the coincidence occurring and surviving all the vetoes.

Now we shall see the general effect of the h veto. Before applying any vetoes, there were 83620 coincidences. After the all vetoes but the h veto there were 65317 coincidences remaining. Now with the above choice of false dismissal probability, viz. 10^{-3}, the h veto reduces the number of surviving coincidences to 57802; i.e. the h veto has removed 12% of the coincidences with only a 0.1% chance of removing any real strong gravitational waves which may have been present. In Chapter 8, we shall see that the h veto is especially good at removing coincidences which are very high in signal-to-noise in one detector, while very low in the other. Intuitively, one would expect such coincidences to be spurious. From an aesthetic point of view, this considerably "tidies up" the coincidence list.

In total, the vetoes used have removed 31% of the coincidences in the original coincidence list. This was achieved with a very minimal probability (a few percent) of removing any real coincidences which may have occurred during normal operation of the detector. A more visual presentation of these results will be given in Chapter 8.

7.3.5 Sensitivity veto (sigma veto)

We already know that there are times when the data are badly non-Gaussian, and I have argued that we should remove data taken from these times. A slightly more subtle point is whether or not we should discard data when the detector is operating with poor sensitivity; i.e. when the standard deviation of the noise is high. I will call this a "sigma veto". This may occur even when the data are well described by a Gaussian noise model.

Clearly, removing data where the noise variance is high in amplitude will improve the sensitivity of the experiment, because the noise remaining will be, *ipso facto*, lower in amplitude. The immediate disadvantage would be that the amount of data remaining would be less, so that one would be setting experimental limits on a shorter and, therefore, less interesting dataset. Notwithstanding, one could argue that this procedure would be worthwhile if (a) the reduction in the dataset were "small", and (b) the limit set by the experiment were to improve by a factor which were "large enough".

I believe that, in principle, such a sigma veto is justified because the stricter the upper limit observed, the better from the point of view of science. However, I have not used such a veto here for two reasons:

- In this coincidence analysis, I have tried, wherever possible, to avoid a posteriori analysis. A sigma veto would by necessity involve rejecting data which we do not like, because those data contaminated the wrong part of our plots. Hence, it is a posteriori analysis. This differs from the other vetoes I have used, because I have implemented them blind, with no explicit control on the upper limit of h.

- Throughout my construction of my vetoing system, I have been aware that this is the first such analysis performed on a long dataset of broadband interferometric gravitational wave data. Consequently, I am loathe to throw away large chunks of data which may contain gravitational wave signals. I also feel that such a choice regarding the quality of the data is too subtle to make in this first attempt at constructing such an analysis system.

I certainly do not object to the use, in principle, of a sigma veto. If the motive for its use is stated explicitly, i.e. that one desires to lower the upper limit on h, then I see no harm in it. However, this was not a major explicit task in my analysis. In any case, it is likely, that such a sigma veto will be used in the near future, in papers produced by our group on the broadband and low pass filtered data; and preliminary analysis (Nicholson 1993) suggests that the improvement in sensitivity will be good for removing a relatively small subset of data.

7.3.6 Event sign veto

There is one veto which we have not yet addressed. This pertains to the fact that in the recording of the events from the original tapes, the modulus of the amplitude of the peak value was taken. In the event lists and coincidence lists, therefore, there is no way to recover the sign of the peak. Although the situation is complicated by complex event morphology such as oscillatory events which cross both the positive and negative threshold, in the simple case of events which are exclusively positive or exclusively negative, it would be quite trivial to veto those coincidences which are of opposing sign. One would need, of course, to return to the original tape and search for each event individually, which would be very time consuming. For future such analysis systems, I recommend that the sign of each peak, as well as perhaps more detail of each event of length greater than one datapoint, should be recorded in the event lists. If this were done, this veto alone could reduce the number of spurious coincidences by up to 50%.

7.4 Interrelations of the vetoes

Finally, we shall look at the interrelations between the vetoes at the levels chosen, as this may tell us something about the way the vetoes work, and may also tell us about physical relations between the housekeeping streams. Potentially, it could also tell us which vetoes are redundant, if any. In this discussion, we shall not include the amplitude veto, which is qualitatively different to the other vetoes in that it uses information from both Glasgow and Munich data together.

7.4.1 Interrelations of Glasgow vetoes

We shall begin with Table 7.6, which shows the numbers of coincidences vetoed by both of each pair of housekeeping streams used from the Glasgow data, i.e. the overlap between vetoes. Although this table appears complicated at first, much can be gained from its patient inspection.

In Table 7.6, the following abbreviations have been used:

Ev Prier Event Primary Error Point signal

Grp Prier Group Primary Error Point signal

Grp Mic Group Microphone signal

Sec Vis Group Maximum Secondary Visibility signal

Grp Seis Group Seismic signal

Gaus Par Gaussian Parameter

The bold numbers down the diagonal are the numbers of the 83620 coincidences vetoed by each of the housekeeping streams, taken from column 2 of Table 7.4. By reading along the rows and columns of the table, one can see the relationship between each pair of housekeeping streams. In the upper right section of the table, the large numbers are the numbers of the events which are vetoed by *both* of each pair of housekeeping streams; while the small numbers in brackets are the number we would expect to be vetoed by chance by both of each pair of housekeeping streams, if there were no physical relationship between them. This number is given by the following calculation:

Let $E_1(t)$ and $E_2(t)$ be sequences of coincident events vetoed by two independent housekeeping streams, forming sets E_1 and E_2, and let N_1 and N_2 be the numbers of coincidences so vetoed. Assume also that the sequence of events vetoed by each housekeeping stream forms a Poisson process, i.e. there is no relationship in the time between each pair of events vetoed, except for a mean arrival delay. This will not be completely true, since the mean time between coincidences is around $(100 \times 3600/83620 \sim)$ 4 s, but it will give us a number to work from.

Table 7.6: Overlap of Glasgow vetoes

Signal	Ev Prier	Grp Prier	Grp Mic	Sec Vis	Grp Seis	Gaus Par
Ev Prier	**238**	11 (15)	0 (0.2)	73 (49)	0 (1.5)	44 (21.4)
Grp Prier	4.6% of Ev Prier 0.2% of Grp Prier	**5133**	0 (3.4)	4429 (1056)	4 (31.4)	2456 (462)
Grp Mic	0.0% of Ev Prier 0.0% of Grp Mic	0.0% of Grp Prier 0.0% of Grp Mic	**56**	50 (11.5)	50 (0.34)	7 (5.0)
Sec Vis	30.7% of Ev Prier 0.4% of Sec Vis	86.3% of Grp Prier 25.7% of Sec Vis	89.3% of Grp Mic 0.3% of Sec Vis	**17206**	511 (105)	7103 (1551)
Grp Seis	0.0% of Ev Prier 0.0% of Grp Seis	0.1% of Grp Prier 0.8% of Grp Seis	89.3% of Grp Mic 9.8% of Grp Seis	3.0% of Sec Vis 100.0% of Grp Seis	**511**	197 (46.0)
Gaus Par	18.5% of Ev Prier 0.6% of Gaus Par	47.8% of Grp Prier 32.6% of Gaus Par	12.5% of Grp Mic 0.1% of Gaus Par	41.3% of Sec Vis 94.3% of Gaus Par	38.6% of Grp Seis 2.6% of Gaus Par	**7536**

Then by the same definition, $[E_1 \cap E_2](t)$ is also a Poisson process; and it has mean and variance given by

$$N_{E_1 \cap E_2} = \frac{N_1 \times N_2}{83620} \tag{7.50}$$

(compare Eq. 2.6).

This is the number given in brackets in each entry of Table 7.6. Thus we can compare the number of vetoed events in overlap between two housekeeping streams with the number we expect by chance.

The percentages in the lower left of Table 7.6 refer to the number of events in overlap as a fraction of the total number vetoed independently by the two housekeeping streams in question. This enables us to say how much one housekeeping stream is doing the job of another.

For example, let us look at the relationship between the Group Primary Error Point stream and the Secondary Visibility stream. Reading along the second row and fourth column, we see that 4429 vetoed events are common to both streams, compared to an expected 1056. A quick mental calculation tells us that the standard deviation is around 30, and so these 4429 common events almost certainly did not occur by chance, i.e. there is a strong physical or causal relationship between the two streams. By reading back down the rows and columns to the bold numbers, we see that 4429 of the 17206 Secondary Visibility vetoed events are also vetoed by the Group Primary Error Point signal, while conversely, 4429 of the 5133 group Primary Error Point vetoed signals are vetoed by the Secondary Visibility signal. The lower left section of the table tells us that this is 86.3% of the Group Primary Error Point vetoes, and 25.7% of the Secondary Visibility vetoes. In other words, most of the job of the Group Primary Error Point veto is done by the Secondary Visibility signal. This strong physical relationship between the two streams is caused primarily by the relationship between the Primary Error Point and the Primary Visibility signal (determined by the behaviour of the primary cavity), and between the Primary Visibility and Secondary Visibility signal (determined by whether the detector cavities are in lock).

Another example is the Secondary Visibility and Group Microphone. Here, nearly all (89.3%) of the Group Microphone signals coincide with a loss of secondary cavity lock. In fact, when the Glasgow detector loses lock, the oscilloscope showing the Secondary Visibility output goes crazy and emits a slight hiss; but if this were responsible for the simultaneous microphone events, one would expect that there should be a microphone event each time there is a Secondary Visibility event, and there should be at least 17206 microphone events. Since this is not the case, the hiss must not be causing microphone events (a possible explanation for this is that our threshold level on the Group Microphone veto was set much too high). Alternatively, the correlation may have been caused by some of the Glasgow experimenters sometimes, but not always, audibly expressing their irritation as the detector lost resonance. In this case, any noises made would have to be within 1 s or so of the loss of lock, in order to fit into the same group. A third possible explanation is that sporadic noises in the laboratory were being picked up by the detector and causing it to lose lock.

Note the low number vetoed by both Group Primary Error Point and Group Seismic signals, compared to the number expected by chance (4 compared to 31.4). After many checks, this discrepancy remained. I can offer no explanation for this apparent anti-correlation.

Note also that there is no unusual correlation between the group and event Seismic events, i.e. there appear to be two types of seismic events revealed: those that last for of the order of one data point[7], and those that last for the order of one group. These types of seismic events appear not to be causally related.

There is much to be learned from Table 7.6.

7.4.2 Interrelations of Munich vetoes

Since we have not used many Munich vetoes, this subsection is going to be short. For completeness, I have included the Munich analogue of Table 7.6, i.e. Table 7.7; but it is not very revealing.

In Table 7.7, I have used these abbreviations:

Grp Tab Group Table Seismometer signal

Grp Low Group Low Frequency signal

Gaus Par Gaussian Parameter

The most noteable correlation here is that nearly all of the Group Table events are also vetoed by the Gaussian Parameter; i.e. when the Group Table seismometer is excited, the data are usually very non-Gaussian. This is gratifying, though not very surprising.

[7] In fact, the seismic data were recorded at 1667 Hz, or once every 6×10^{-4} s

Table 7.7: Overlap of Munich vetoes

Signal	Grp Tab	Grp Low	Gaus Par
Grp Tab	**123**	0 (0.01)	107 (0.5)
Grp Low	0.0% of Grp Tab 0.0% of Grp Low	**12**	0 (0.05)
Gaus Par	87% of Grp Tab 31.4% of Gaus Par	0% of Grp Low 0% of Gaus Par	**340**

Since these data have not been shown in time series form before, as the Glasgow output streams have (Watkins 1991), I have included a sample output here, in Fig. 7.22.

The data shown in Fig. 7.22 are from experimental times 57h 20m (2.064×10^5 s) to 62h 34m 45s (2.25285×10^5 s). This period was chosen randomly in the sense that these data happened to be on the hard disk awaiting separate analysis, at the time I wanted such a printout. This figure depicts the Gaussian Parameter, the Detector Output Group Sigma (the standard deviation of each group of data after weighting and calibration, so it is essentially the inverse of the standard deviation of the raw data), the Table Seismometer signal, the Suspension Seismometer signal, and the Low Frequency signal. They are the group data only.

One can see immediately that there are a few very strong events which appear in all the data streams in most cases. These occur at 2.064×10^5 s (against the left edge of each graph), 2.117×10^5 s, 2.170×10^5 s, 2.232×10^5 s, and 2.234×10^5 s experiment time. The event at 2.117×10^5 s is notable by its absence from the Suspension Seismometer signal, while those at 2.117×10^5 s, 2.232×10^5 s and 2.234×10^5 s are curiously missing from the Low Frequency signal.

In addition, if one looks carefully at the Gaussian Parameter stream, especially between 2.180×10^5 s and 2.240×10^5 s, one can see a regular series of very small spikes, at a Gaussian Parameter value of around 1.3. The time between these events is around 500 s=8 minutes

20 seconds, so it is very likely that these spikes coincide with water cooler events. It is interesting that most of these events are single, where we expected the water cooling events to be double. The most likely explanation for this is that of the opening and closure of the water valve, one or the other had a much greater effect on the noise. In fact, some of the events which appear to be part of this pattern, particularly those between 2.118×10^5 s and 2.144×10^5 s, seem to be double, with the two events separated by about one quarter of a time division, or about 50 s.

It is compelling that these patterns are real because there are very few spikes, even at this small amplitude, which do not fall into a regular 500 s cycle.

There is also evidence that such a pattern is contained in the Detector Output Group Sigma. However, the events are at lower signal-to-noise and are harder to see. They can be seen fairly well, though, between 2.118×10^5 s and 2.135×10^5 s, where they appear to be in coincidence with the events in the Gaussian Parameter data, and where the double nature of two of the events is clearer than in the Gaussian Parameter data.

In all five data streams, one can see the tape change hiatuses, occurring at 2.100×10^5 s, 2.139×10^5 s, 2.177×10^5 s, and 2.216×10^5 s.

The five large spikes remain unexplained, as they are not part of the water valve pattern. There are two remaining predicted populations of strong events which could be visible in these data, being the laser argon refill events, and other causes of loss of lock. There were 24 laser refill events recorded during the experiment[8], each a pair of "clicks" separated by 126 ± 0.5 s, with the pairs being separated by about 200 minutes from each other, each of which events caused a loss of lock (Alarm) (Rüdiger 1990). There were a total of 78 losses of lock during the 100 hours, leaving 30 losses of lock not caused by laser refill. In total, these losses of lock occur at something like less than one event per hour, so the events seen in Fig. 7.22 are occurring at something like the correct occurrence rate.

The double event at 2.232×10^5 s is probably an argon laser refill event since the two events of the pair are spaced about 130 s apart, and since there should be at least one refill-caused double event during the five hours or so of these data (since they occur fairly regularly about once every 200 minutes), and since it occurs about 62h experiment time, which corresponds to a refill event recorded by the experimenters (Rüdiger 1990). There

[8]There may also have been a few laser refills which occurred while the tape recorder was switched off, as the alarm is deactivated during these times.

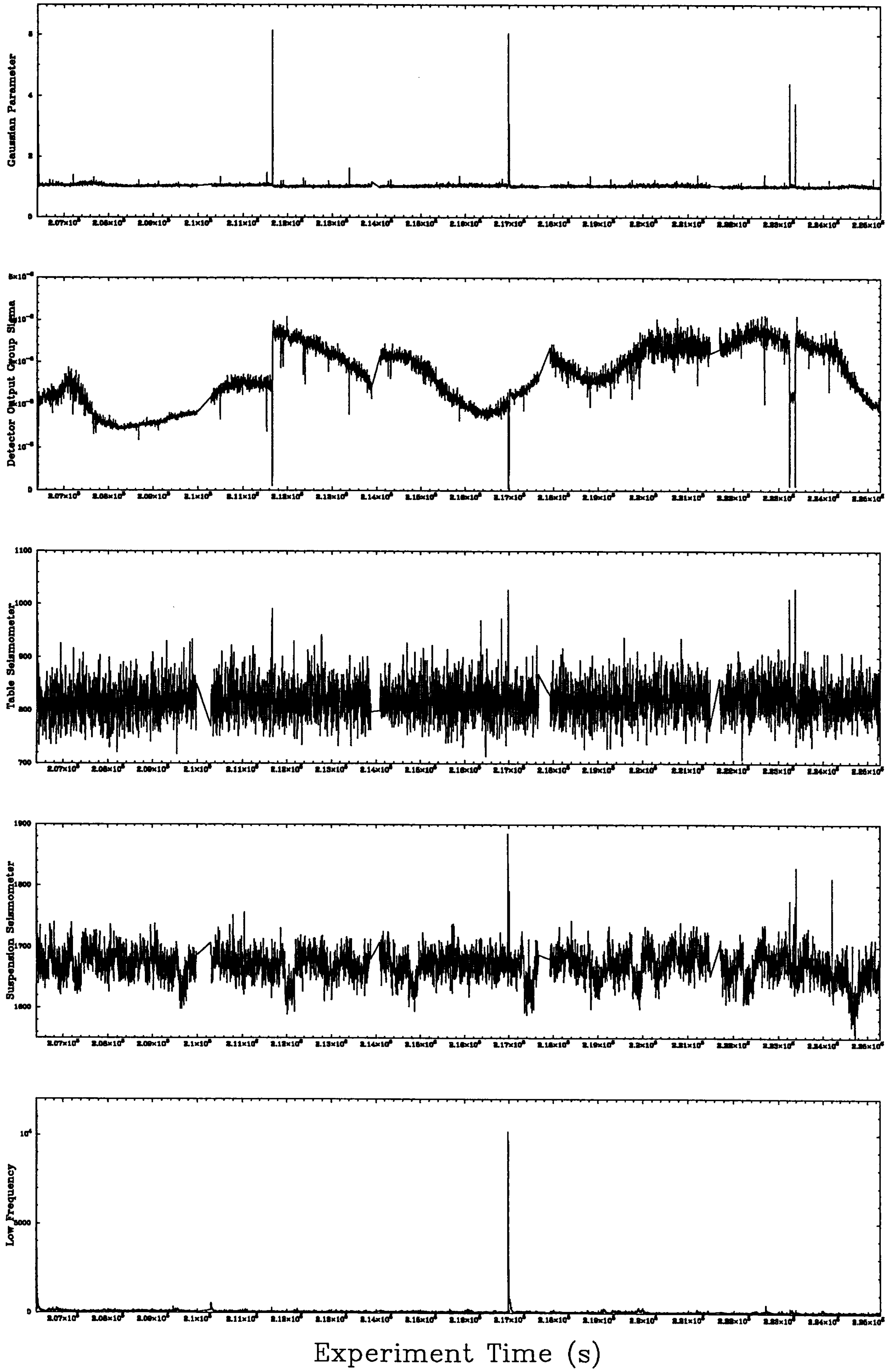


Figure 7.22: Munich group data (57h 20m to 62h 34m)

may also have been a laser event 200 minutes previously, which has been missed during the tape change at 2.102×10^5; or it could be that the previous event occurred just off the left edge of the plot.

This leaves three peaks which are probably losses of lock not caused by laser refill events. It is not clear why some of these events do not appear in certain group data while others do, and I shall not speculate here on why this should be so.

In any case, this discussion of very high peaks is probably irrelevant from the point of view of vetoes, because if they are losses of lock, they will not have been recorded by Nicholson's event program, and so will not appear in the coincidence list.

This concludes our discussion on vetoes. The results of the coincidence experiment, including the effects which these vetoes have on the coincidences, are presented in the next chapter.

Chapter 8

Results of the Coincidence Analysis

Having outlined the coincidence problem (Chapter 5); stated my analysis method, considered the individual problems of the Glasgow and Garching datasets, and run the coincidence program (Chapter 6); then done what is possible to remove untrustworthy sections of data (Chapter 7); I am now in a position to present the results. These include: a straight presentation of the coincidence list, in the signal-to-noise plane of the two detectors, including the effect on this of the vetoes; individual histograms of signal to noise of the events in the independent outputs of the two detectors, which have not been presented before; the a priori probabilities of the least likely coincidences found; and the upper limit on h, the broadband gravitational wave strain amplitude, for the experiment.

8.1 Coincidence list

The output of the coincidence program took the form of a list of pairs of events, one event from each of the detectors, where a coincidence is defined as in Definition 6.1. Each entry in the list contained as much of the relevant information about the individual events as was recorded by the individual analysis programs, such as the times of the events, the tape numbers, the peak signal-to-noise values for the events, and all the housekeeping data denoted "retained" in Tables 7.1 and 7.2.

After removing those coincidences contained in repeated blocks (see Section 6.3), there remained a list of 83620 coincidences. This compares fairly well with the number expected from the Poisson calculation in Section 6.1.1, viz. 1.04×10^5. This discrepancy will probably have arisen from a combination of the following factors:

- there are many short gaps and a few very long gaps in the data of both detectors (e.g. tape stoppages, tape jams, Alarm signal in Munich) which will render Eq. 6.1 at best a crude approximation, since this equation presupposes that the noise is stationary;
- there are also nonstationarities in the *quality* of the data, with the variance of the noise commonly changing by factors of two in both detectors; as well as changes in the distribution of the noise, which is not always, if ever, completely white and Gaussian;
- another supposition of Eq. 6.1 is that the individual data streams were Poissonian. However, the fact that many of the events in both detectors are more than one datapoint long shows that this assumption is not correct. Further, the event lists are lists of contiguous sets of datapoints, all of which datapoints cross threshold, whereas Eq. 6.1 really applies to lists of individual datapoints which cross threshold irrespective of the behaviour of neighbouring datapoints (also presupposing that neighbouring points are independent).

Figure 8.1 shows how these 83620 coincidences were distributed in time. One sees immediately that the numbers of coincidences are not spread uniformly in time, reaffirming the hypothesis that the data are nonstationary. Note particularly the gap between 64h and 66h E.T., caused by a problem with the tape deck at Garching (Rüdiger 1990), and other low or empty bins caused by tape problems: either at the sites (as reported in the experimental logs recorded by both teams), or at the analysis stage in Cardiff (Watkins 1991; Nicholson 1991).

In Fig. 8.2, I show the distribution of time delays between the peak times of each event of each coincidence. The distribution is very flat, as expected. Note that there are not obviously more coincidences between -4.6 ms and $+4.6$ ms than outside these values, so at least there is no huge population of real, strong, broadband gravitational waves arriving at the detectors. This plot doesn't really tell us anything new, but it is at least consistent with the models we have.

We can now compare the behaviour of the two detectors in signal-to-noise space: this will lead later to a calculation of the a priori probability of the coincidences. Figures 8.3, 8.4, and 8.5 show the coincidences plotted in the signal-to-noise plane of the two detectors,

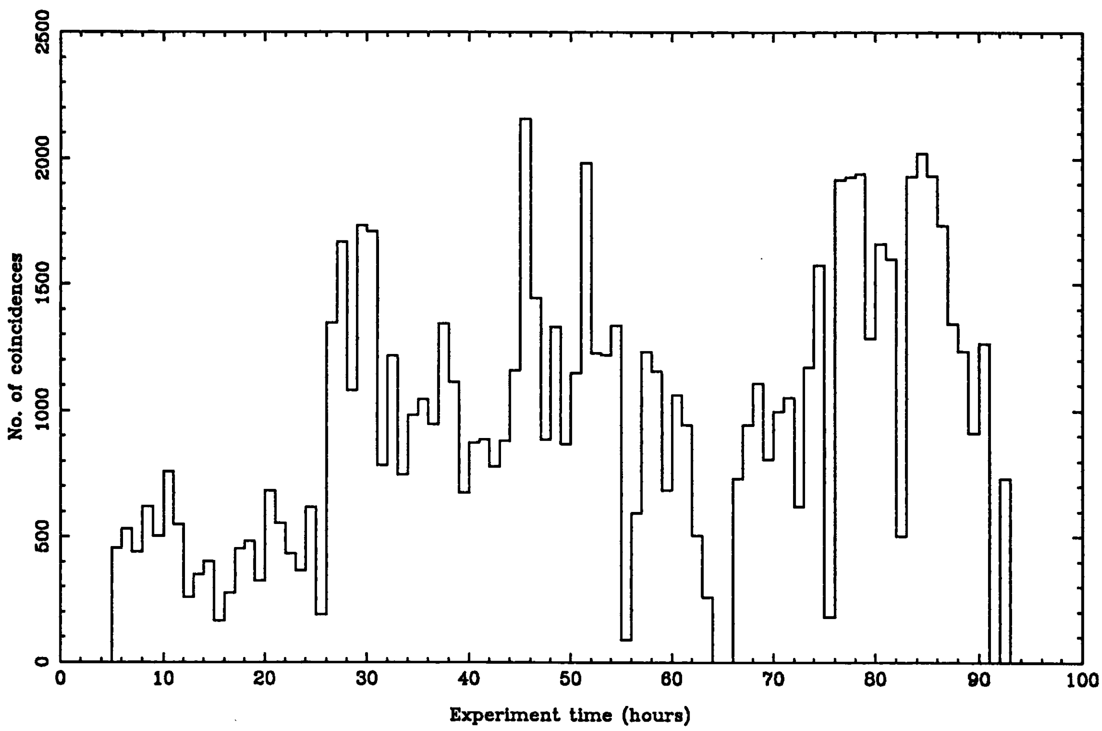


Figure 8.1: Numbers of coincidences in each hour (5h–93h, no vetoes)

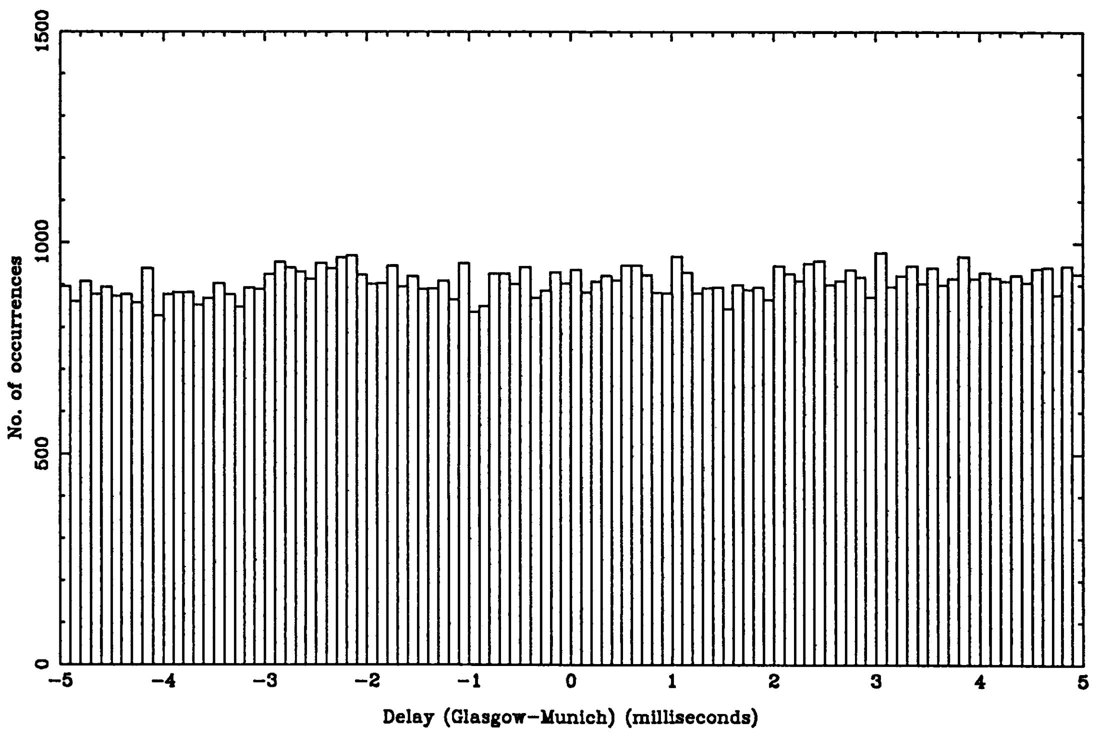


Figure 8.2: Numbers of coincidences at each delay between observations (5h–93h, no vetoes)

respectively after no vetoes, after all the vetoes except the h veto, and after all the vetoes including the h veto. The origin of all three plots is $S/N = 4.0$, i.e. the threshold for both detectors. In Fig. 8.3, one sees that there is a strong high amplitude tail in both detectors: if the experimental data were Gaussian, we would expect the highest signal-to-noise value seen to be around 6.4. In fact, the highest seen are around 37 (Glasgow) and 71 (Garching). This behaviour is greatly improved by the vetoes: while the "standard" vetoes considerably thin out the low S/N part of the plot, the h veto is particularly effective at removing those coincidences along the axes, i.e. those with high S/N in one detector and low S/N in the other. This is as expected. The h veto is particularly effective at tidying up the Munich axis, perhaps because the Munich detector was often slightly the less sensitive of the two detectors (the standard deviation of the noise in Munich was typically higher, in strain units, than in Glasgow), resulting in the inferred gravitational wave amplitude being much higher in Munich than in Glasgow for these coincidences, and hence being vetoed.

Overall, the numbers of coincidences were 83620 (no vetoes), 65317 (all except h veto), and 57802 (all vetoes). Thus, the vetoes have performed well in tidying up the coincidence list. By implementing an "on line" system, as I have advocated in Chapter 7, I believe this could be improved again, other things being equal. In addition, one could make an argument that one should keep only those data where the detectors are not only performing satisfactorily, but are performing well, with close to optimum sensitivity (a "sigma-veto": see Section 7.3.5). This would probably involve rejection of much more data, but with the advantage of a great improvement in the experimental limits set. I shall not do this here, but this will be used in any future research papers based on these data (e.g. Hough *et al.* 1993).

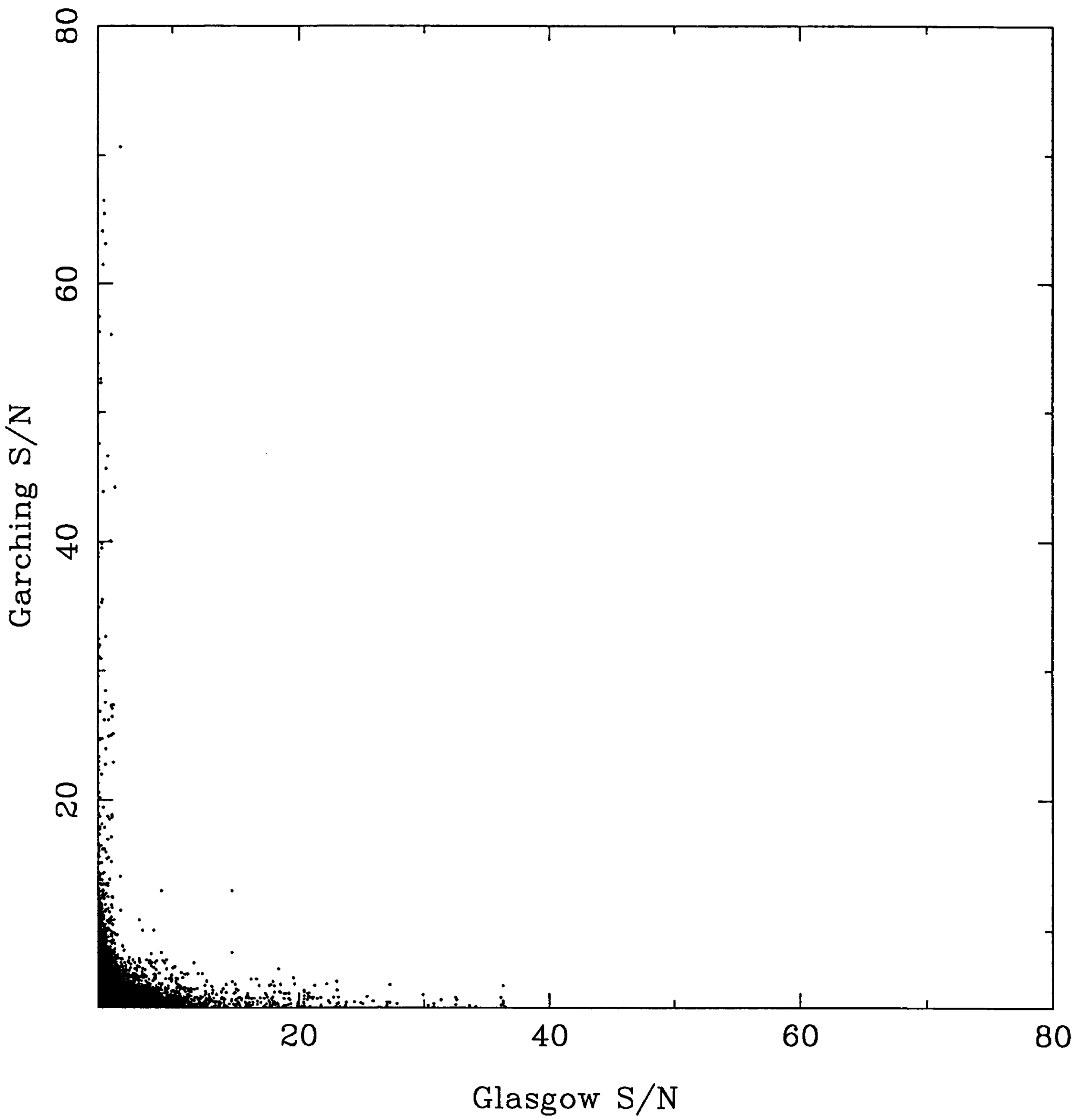


Figure 8.3: Scatter diagram of the signal-to-noise of the coincidences (5h–93h, no vetoes)

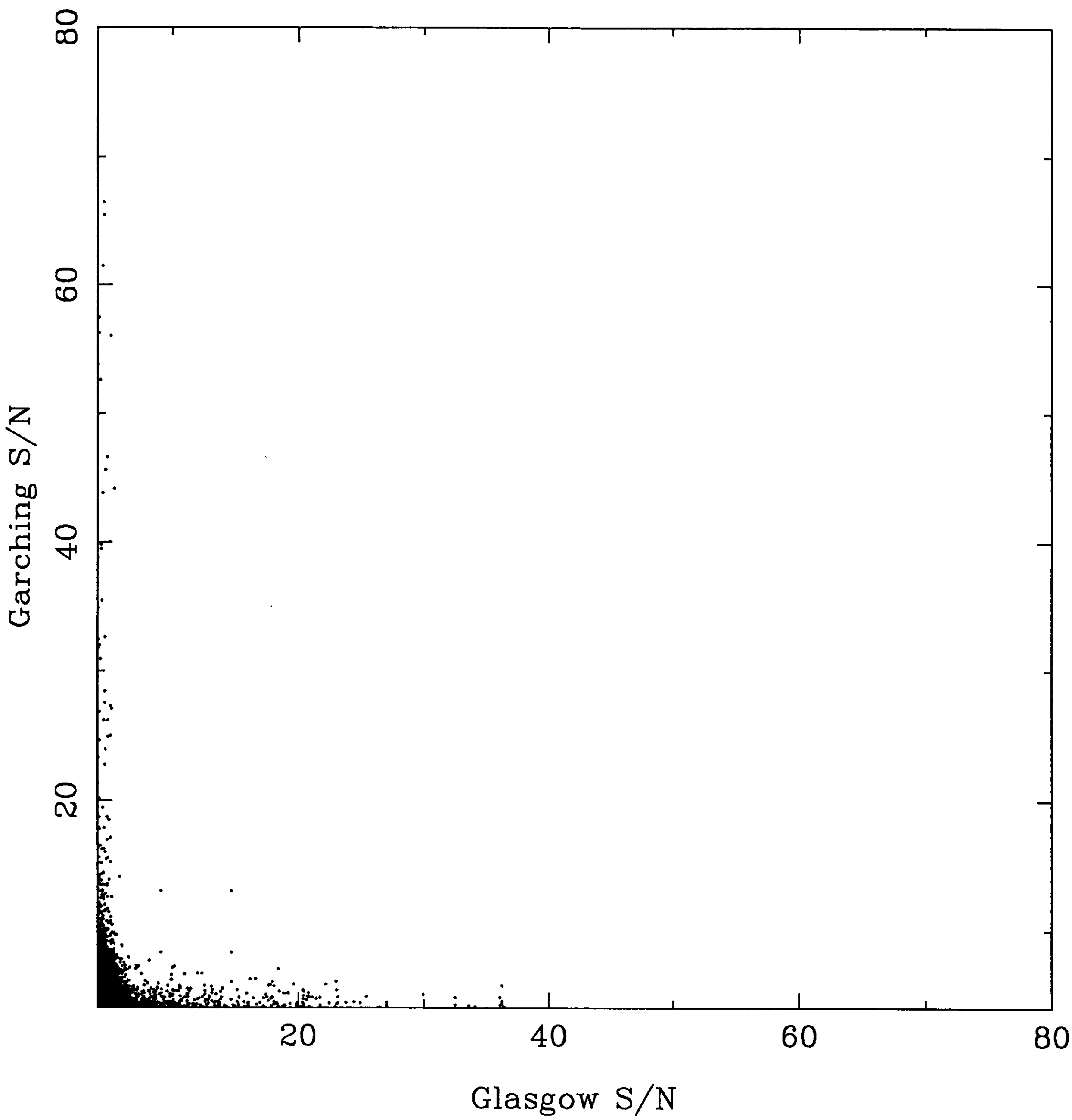


Figure 8.4: Scatter diagram of the signal-to-noise of the coincidences (5h–93h, all vetoes except h-veto)

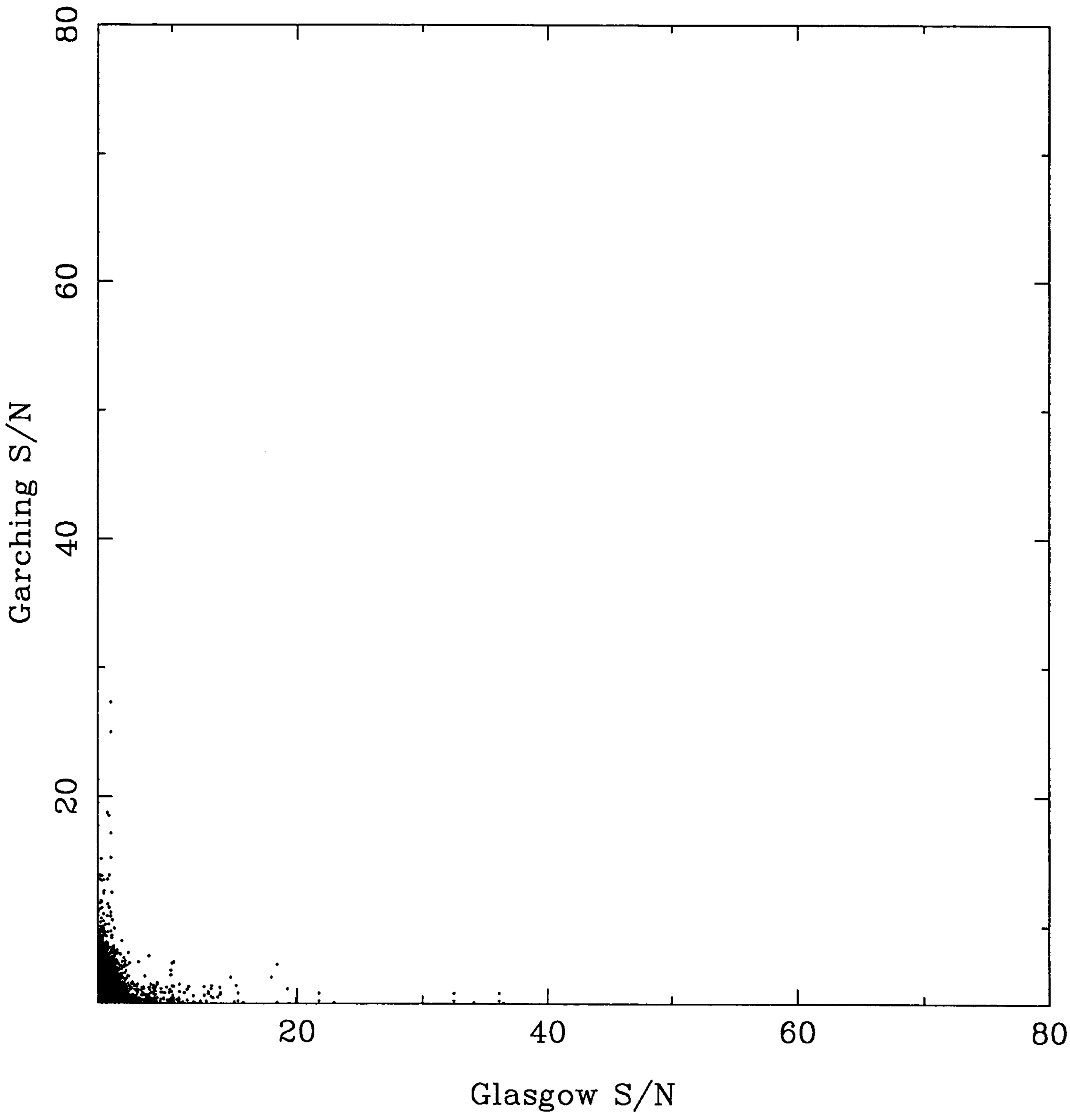


Figure 8.5: Scatter diagram of the signal-to-noise of the coincidences (5h–93h, all vetoes)

8.2 Probability analysis

8.2.1 Method of probability calculation

Having the signal-to-noise values of the coincidences enables us to calculate the a priori probability of the coincidences, based on the empirical distributions of the individual data streams, and based on the hypothesis that the two noise streams are independent. The individual distributions give us the probability of obtaining an event of a certain signal-to-noise range or higher, and we can then multiply these two probabilities together, also accounting for the width of the delay window, and also including the length of the experiment, to obtain the probability per experiment of obtaining a coincidence of the two signal-to-noise measurements or better.

In doing this, we must be careful to use the same sample of events from which the coincidence list was drawn, i.e. to use those datastreams with the same vetoes applied. This is easy in the case of the independent vetoes (Glasgow housekeeping and Gaussian Parameter, Garching housekeeping and Gaussian Parameter), but is more problematic in the case of the h veto, since those events removed by the h veto are selected in a way which is dependent on both datasets simultaneously. Thus, what we have done to the coincidence dataset cannot exactly be replicated with the individual (independent) event lists.

Now, in very simple terms, all the h veto has done is to remove a random subset of the coincidences; and therefore, the probability of surviving the h veto is the ratio of the number of coincidences surviving the h veto to the number of coincidences before the h veto. This is equal to 57802/65317: I shall call this fraction f_h.

However, it is not that simple: the h veto will preferentially remove those coincidences where the S/N is high in at least one of the events (this is clearly seen from the definition of the h veto, Eq. 7.49). Thus, coincidences with a higher value of h will have a lower probability of surviving the h veto than those with lower h. In fact, for a coincidence with given S/N and σ in both detectors, the h veto is *deterministic*, i.e. it is not random, and the probability of surviving it is either 0 or 1 depending on the measurements. As far as I can see, the only robust empirical way to calculate the real probability of finding a certain coincidence, and its surviving all the vetoes, including the h veto, is to calculate the probability from the "background method", outlined in 6.1.2; but as we saw with this method, there may be even more problems with the interpretation of the probability of the coincidences. As a compromise, we shall adopt the simple approach, that the probability of a coincidence surviving the h veto is the number of coincidences surviving the h veto divided by the number of coincidences tested by the h veto. [1]

Choose a pair of simultaneous points in the two detectors, as shown in Fig. 8.6. (Note: we must be careful here to respect the different sampling rates of the two detectors.)

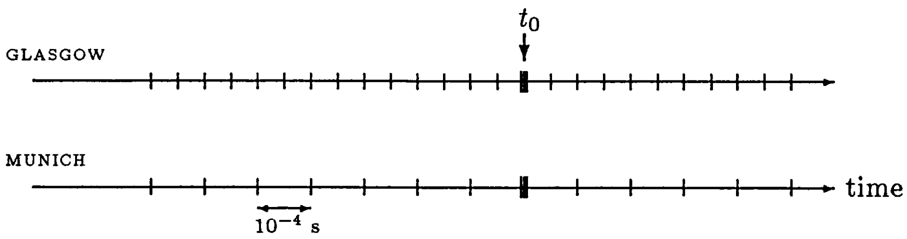


Figure 8.6: Simultaneous events at the two detectors. The normal non-event samples are indicated by thin ticks, while the arrowed events at time t_0 are indicated with thick lines. The probability of these two events occurring at time t_0 is equal to $p_g p_m$ (see text).

Let p_g be the probability that an event of signal-to-noise greater than or equal to a certain value occurs at the indicated Glasgow datapoint or sample; similarly for Munich, for another signal-to-noise value. These probabilities can be empirically derived from the behaviour of the detectors: we need no longer assume that the detectors' outputs are Gaussian in distribution, only that they are white, stationary, and independent of each other. Then the probability of obtaining an event of these signal-to-noise values or higher in both detectors at this time is $p_g p_m$.

Now we allow the time window to vary. For each Glasgow datapoint, the nearest 92 Munich datapoints fall within the time window, and any one or more of them could give the

[1] For anyone who is not satisfied by this, and who believes that the probability we derive for the least probable coincidences could be greatly affected by the fact that these have may have higher h values, I would answer that at the limit on h which we will set, there are four coincidences close by in h space; and since we believe gravitational waves are rare (at this sensitivity), some or all of these will almost certainly be noise-induced. Hence, even if the least likely coincidences are made considerably less likely to survive by their having a high h value, we still find at least two or three random noise-induced coincidences which have high h and survive the h veto. In any case, therefore, the probability per experiment of the least likely coincidences will not be far below 0.5.

required signal-to-noise value or higher. Thus, the probability that the Glasgow datapoint is an event of this signal-to-noise or higher (with probability p_g), *and* that there is at least one Munich event within the time window with greater than or equal to the appropriate signal-to-noise (with probability p_m), is, strictly,

$$p_g \sum_{r=1}^{92} \mathrm{C}_r^{92} p_m^r (1-p_m)^{92-r}; \tag{8.1}$$

i.e. the Binomial distribution, where

$$\mathrm{C}_r^{92} = \frac{92!}{(92-r)!r!}. \tag{8.2}$$

Since p_m is small, however, we will obtain the Poisson approximation with mean $92p_m$; so that the probability of this coincidence is

$$p_g \sum_{r=1}^{92} \frac{(92p_m)^r e^{-92p_m}}{r!} \tag{8.3}$$

$$= p_g e^{-92p_m} \sum_{r=1}^{92} \frac{(92p_m)^r}{r!}. \tag{8.4}$$

Now $92p_m$ is also small, so only the $r = 1$ term contributes, and the probability is, as expected,

$$92p_g p_m. \tag{8.5}$$

Note that the events in Glasgow are typically much longer than one light travel window apart. Consequently, the sets of Munich datapoints in the window of each Glasgow event will not overlap; i.e. all the groups of 92 Munich datapoints under consideration will be independent of each other. Let the number of Glasgow datapoints in the experiment be N_g. Then the probability per experiment of having no such coincidences (i.e. none with these signal-to-noise values or higher) will be the Poisson probability of obtaining 0 events, with expected value $N_g 92 p_g p_m$ — again we are using the Poisson approximation to the Binomial, because $92p_g p_m$ is small and N_g is large — viz.,

$$\mathrm{P}(0) = \frac{(92p_g p_m N_g)^0 e^{-92p_g p_m N_g}}{0!}; \tag{8.6}$$

and so the probability of at least one such coincidence during the experiment is $1 - \mathrm{P}(0)$, i.e.

$$1 - \exp(-92p_g p_m N_g). \tag{8.7}$$

Now we allow for the h-veto. The probability that our original two datapoints have greater than or equal to the required signal-to-noise *and* pass the h veto, is given by

$$p_g p_m f_h, \tag{8.8}$$

where f_h is defined earlier, so that $p_g p_m f_h$ replaces $p_g p_m$ in 8.5. Then the probability of such a coincidence occurring at least once during the experiment and surviving the h-veto, is given empirically by

$$1 - \exp(-92 p_g p_m f_h N_g); \tag{8.9}$$

which we can calculate.[2]

Now, suppose we have a coincidence of signal-to-noise values $(S/N)_G$, $(S/N)_M$ in Glasgow and Munich respectively. Then the probability, p_g, that a given datapoint in the Glasgow stream takes signal-to-noise value $(S/N)_G$ or greater is equal to the sum of the numbers of events in the corresponding bin or higher, divided by the total number of Glasgow datapoints in the experiment. Similarly, one obtains p_m, the probability for Munich. These go into Eq. 8.9 to give the probability, per experiment, of obtaining a coincidence with signal-to-noise values greater than or equal to $(S/N)_G$ and $(S/N)_M$ respectively.

8.2.2 Results of probability calculation

In Figs. 8.7 and 8.8, I show the signal-to-noise distribution of the event lists for Glasgow and Munich, for the whole usable part of the experiment (5h–93h experiment time). The vetoes, except the h veto, have been applied.

The lowest values of signal-to-noise occur at 4σ, the threshold. In both figures, the ordinate axis has been truncated to allow the eye to see what happens at high S/N: the height of the first bin is actually $\sim 10^6$ in both Figs. 8.7 and 8.8.

Then, using the method given in the previous section, the result is that, for all the 57802 coincidences passing the vetoes, the smallest probability of any one coincidence during the experiment is 0.77; i.e. the least likely coincidence, based on the independent distributions of the two detectors, occurred in this experiment with probability 0.77. On the basis of this, there is nothing to suggest that anything unusual has occurred during the experiment.

[2]Note that, if we had done the analysis from the point of view of Munich, 92 and N_g would have been replaced by 184 and N_m; N_m being the number of Munich datapoints in the experiment, which is 1/2 of N_g. Therefore, Eq. 8.9 would not, of course, be affected.

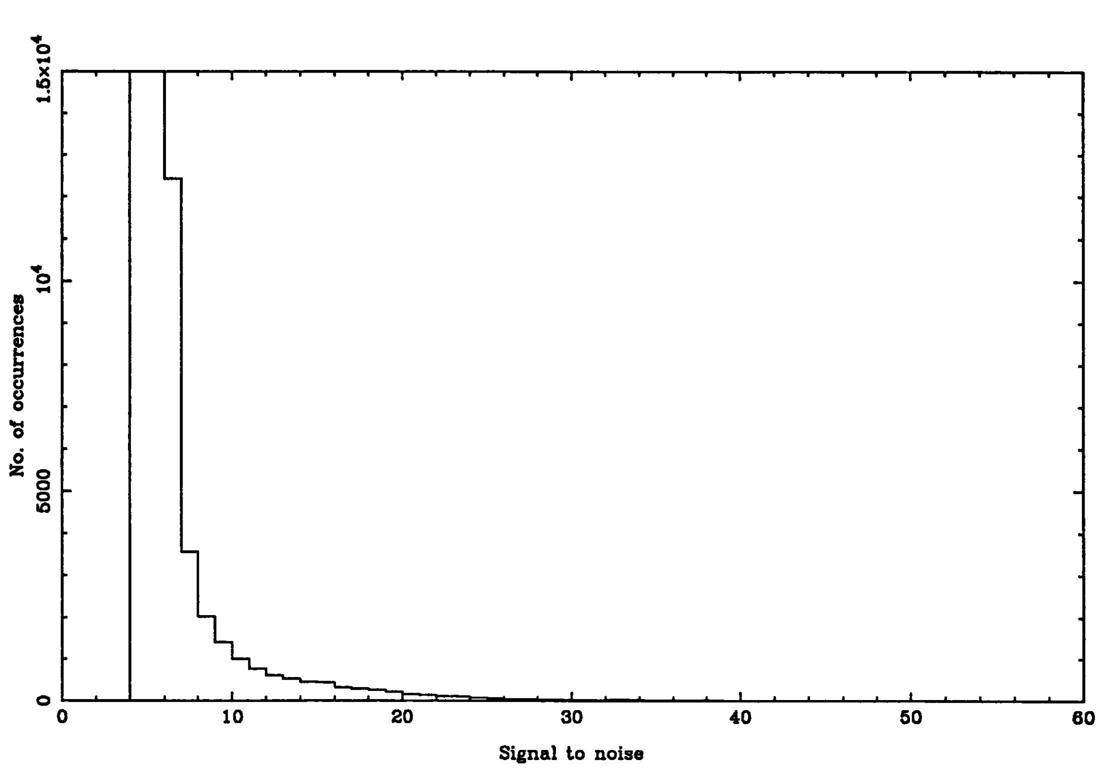


Figure 8.7: Distribution of signal-to-noise of Glasgow events (event list, 5h–93h, all Glasgow vetoes)

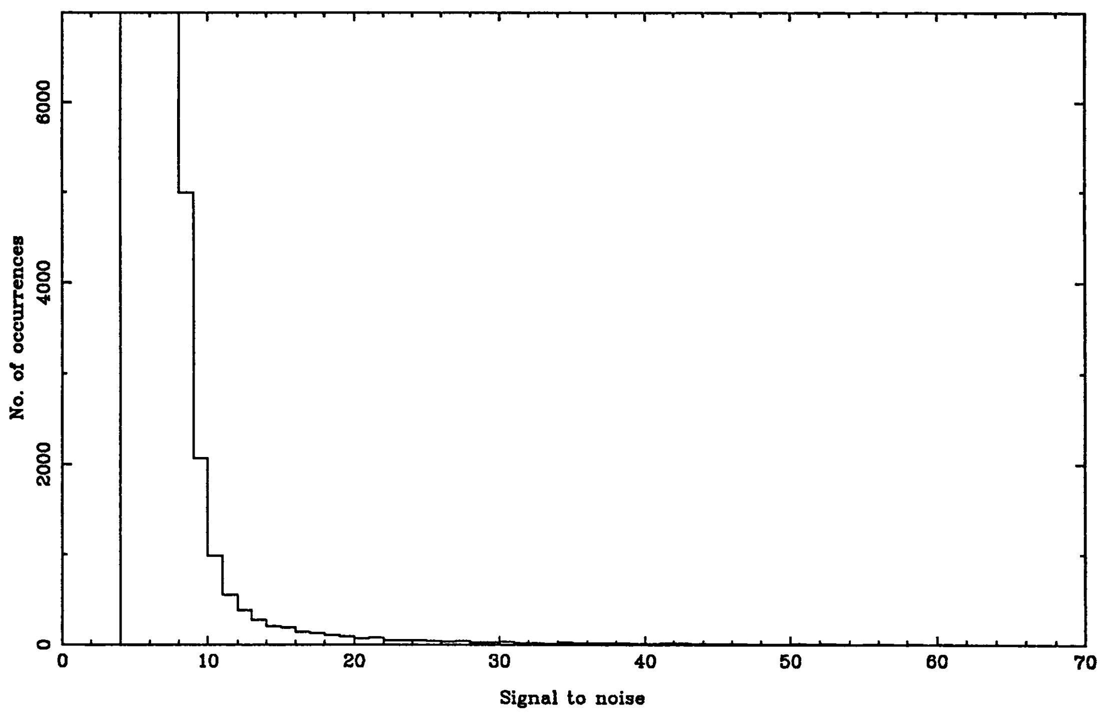


Figure 8.8: Distribution of signal-to-noise of Munich events (event list, 5h–93h, all Munich vetoes)

The full distribution of coincidence probabilities is given in Fig. 8.9. Again, the ordinate

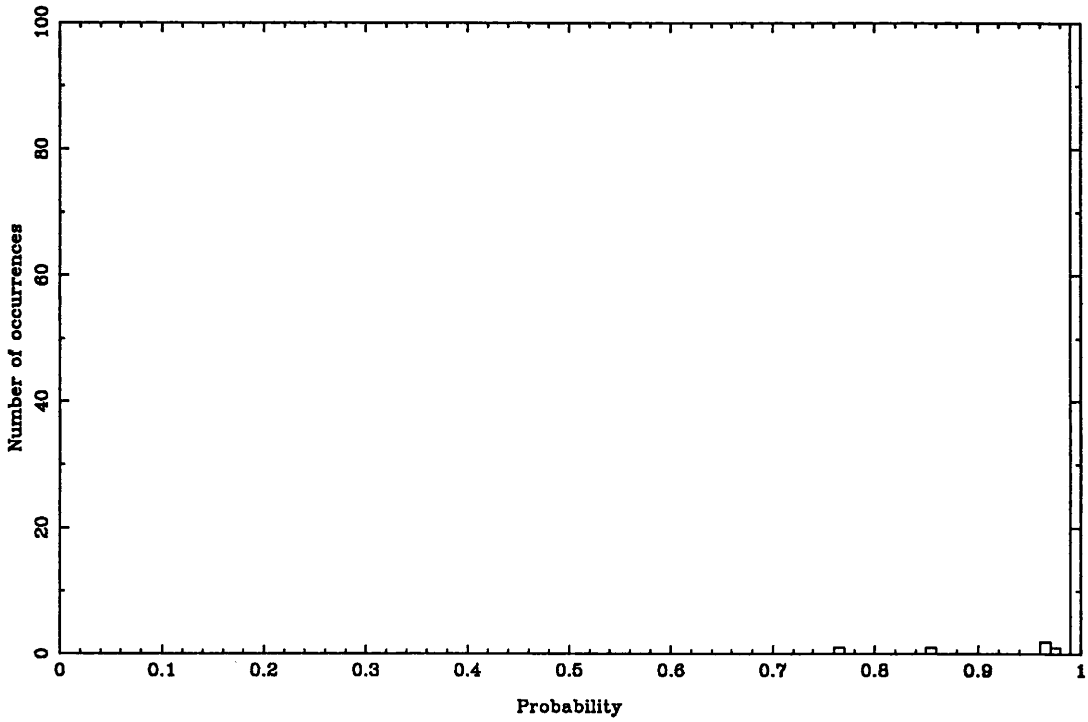


Figure 8.9: Distribution of empirical probabilities of coincidences (coincidence list, 5h–93h, all vetoes)

axis has been truncated. As expected, most of the coincidences occurred with probability $\gtrsim 0.999$, and only a few have probability between this and 0.77. Thus, again, there is nothing to suggest any very unusual coincidences, and so nothing to challenge the null hypothesis that the event lists and coincidence list is composed purely of noise events. Of course, the result is also consistent with there being real gravitational waves present, but any gravitational wave experiment is consistent with this hypothesis: the null hypothesis is the one we accept unless there is reason to doubt it.

8.3 Upper limit on h

Our final result is the astrophysical upper limit on h, the largest value of the dimensionless broadband strain recorded during the experiment, subject to its passing the vetoes and any other restrictions which we have placed on the data. Of course, each coincidence has associated with it two values of h, one each for the two detectors, so we need a way of combining the two values. A complicated way of doing this would be to weight the two observations according to the sensitivity of the detectors — more sensitive meaning more weight. However, this is slightly arbitrary. Moreover, since the detectors were, by and large, of the same sensitivity to within a factor of two, for most of the experiment, it is also unnecessary. For our purposes, it is sufficient simply to average the two observations, i.e. to calculate $(x_g + x_m)/2$, where x_g and x_m are the two measurements.

With this method, it is found that the upper limit on h for the experiment is

$$h = (6.8 \pm 1.3) \times 10^{-16}. \tag{8.10}$$

The measured value is the average of peak values 2.7×10^{-16} and 1.1×10^{-15} in Glasgow and Munich respectively. The experimental error is the combined error of the two observations[3]. Although the observed values are different, this coincidence is tolerated by the h veto because a not unreasonable noise contribution in the detectors could have shifted the same real gravitational wave amplitude to these values.

In Fig. 8.10 I show the h values for all the coincidences in the experiment which pass the vetoes. For brevity, I have not included the same plots without the vetoes.

One can see, firstly, that there is a huge bump in the high h region of the histogram, almost certainly due to some dichotomous behaviour in the sensitivity of one or both of the detectors. In future such experiments, this could be removed outright, as discussed earlier. This could greatly lower the upper limit, and thus improve the sensitivity of the experiment. One also notices the rather straggly high amplitude tail, perhaps due to the strong high amplitude part of the individual S/N distributions of the two detectors.

For comparison, the limit on h without the h veto was 1.7×10^{-15}, while the limit with no vetoes at all was 4.1×10^{-15}. Thus it can be seen that even with this admittedly crude, and perhaps unnecessarily tolerant system of vetoing, the sensitivity of the experiment has been improved dramatically (almost by a factor of 10) by its use.

This concludes the results of the broadband coincidence analysis of the Glasgow–Munich coincident 100 Hour Experiment.

[3]This comes from the well-known property of Gaussian distributions that if two random variables x_g and x_m have mean 0 and variances σ_g^2 and σ_m^2, then the random variable $(x_g + x_m)/2$ has standard deviation $\frac{1}{2}(\sigma_g^2 + \sigma_m^2)^{\frac{1}{2}}$.

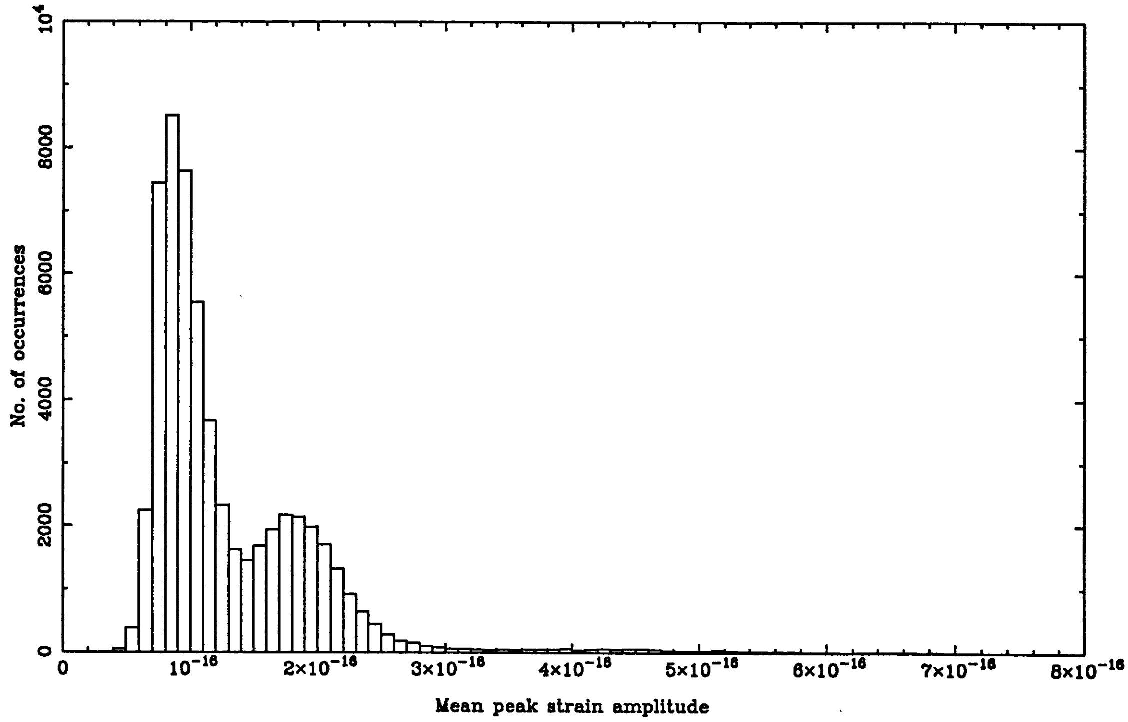


Figure 8.10: Distribution of peak h values for coincidences, averaged over the two detectors (5h–93h, all vetoes)

Chapter 9

Conclusions of the Coincidence Analysis

9.1 Astrophysical conclusions

We have two main conclusions from the point of view of astrophysics.

a) The probability of the coincidences occurring by chance, based on the empirical distributions of the individual detectors, and based on the assumptions that the noise outputs of the two detectors are independent, white and stationary, are as given in Chapter 8. The lowest probability observed is about 0.77. On the basis of this analysis, there is nothing to suggest that any unusual coincidences have been seen. This probability would have to have been $\lesssim 4 \times 10^{-4}$ to challenge the null hypothesis. Of course, our assumptions of whiteness and stationarity of the noise are not correct, as I have demonstrated. However, I would guess that the extent to which these assumptions are wrong is not enough to affect dramatically the result one way or the other — certainly not enough to lower the smallest probability to the 10^{-4} region.

b) The upper limit of the broadband gravitational wave strain amplitude for the experiment is 6.8×10^{-16}, as stated in Chapter 8. This is the first upper limit set on 10 kHz broadband gravitational wave signals. The vetoes were effective at removing many coincidences which occurred in data which should not be regarded as trustworthy; and they were also effective at lowering the upper limit on gravitational waves. For clarity, this information is given here again, in table form (Table 9.1).

Table 9.1: Upper limit on h for various vetoing systems

Vetoes used	Upper limit on h	No. of coincidences
All vetoes	6.8×10^{-16}	57802
All except h-veto	1.7×10^{-15}	65317
No vetoes	4.1×10^{-15}	83620

In particular, the numbers quoted for "No vetoes" are not to be taken seriously, as these include data recorded while the Glasgow interferometer cavities were off-resonance.

9.2 Experimental conclusions

From the point of view of the experiment, and what can be learned for the future of gravitational wave data analysis, there are many more conclusions. In the course of this research, in writing this thesis, I have achieved the following.

1. Written a working program which synchronises the two lists of events from the broad-band time series data, and compares them to look for coincidences in the relevant time window. I have also solved this problem, in the case of a serial code, for simultaneous event lists from low-pass and coalescing binary filters. I have indicated how this could be modified for the case of a parallel code.

2. Provided a list of coincidences from the experiment, and displayed information therefrom in ways which are useful and revealing.

3. Investigated the housekeeping data for use as vetoes against untrustworthy data, and employed them as such, and investigated their interdependence. In doing so, I have greatly improved the sensitivity of the experiment, and improved the behaviour of the noise, while sacrificing only one percent of the useful (on resonance) data.

4. Created and employed a veto based on the two observed amplitudes of a coincidence, which again has greatly improved the upper limit of the experiment. This was with a minimal risk of removing any real gravitational wave signals.

5. Analysed the gravitational wave output of both detectors, with special regard to non-Gaussian noise contributions, and binned their frequency versus strength (the second

of the four *Garching objectives* cited in Section 5.1.1).

6. Rehearsed logistics of data exchange, archiving and retrieval: much has been learned about the problems of handling and storing large datasets, and using Exabyte tape drives. (This was the third of the four *Garching objectives.*)

7. Searched for gravitational wave signals in the data and found that, based on the null hypothesis, there is no evidence to suggest or force us to conclude that gravitational waves have been detected during the 100 Hour Experiment. (This was the fourth and last of the *Garching objectives.*)

8. Set the first upper limit on 10 kHz broadband gravitational wave signals.

9. Demonstrated that the general analysis methodology — that of setting thresholds on the individual datasets, searching through the lists of threshold-crossing events for coincidences, removing untrustworthy events, and assigning a probability to each of the coincidences based on the individual detector output distributions — is feasible for broadband interferometric data, and gives reasonable results.

10. Made a list of comments and suggestions, for the interest of future experimenters and analysis teams, based on my analysis and experience with the 100 Hour Experiment. This is given in Appendix B.

11. Finally, I have laid the foundations of an automated system for coincidence analysis of interferometric gravitational wave data.

It is to be hoped that, through the brilliance and hard work of my many colleagues throughout the world, as well as perhaps through the small contribution contained in this thesis, the detection of gravitational waves will become a reality in the near future. Soon after, I hope, an observational science of Gravitational Wave Astronomy will join with its Electromagnetic, Particle, and Meteoritic counterparts, in opening a new window on the universe.

Appendix A

Listing of Coincidence Program (Prototype Version)

I include here a listing of my FORTRAN program, which was the first working coincidence program for the cross-comparison of event lists from laser interferometer gravitational wave data. It was later amended and simplified by Nicholson and me. For brevity, I have not listed the later version, which was the one which was actually used to generate the results contained in this thesis.

```
      PROGRAM TESTCC

******PROGRAM TO SEARCH FOR COINCIDENCES BETWEEN GLASGOW AND MUNICH
******PROTOTYPE DETECTORS. CALLING SUBROUTINES TESTLGB,TESTLM,GTABLE,
******MTABLE AND FUNCTIONS GEVLOC AND MEVLOC.
      INTEGER GTAPNO,MTAPNO,GEV,MEV,GPOINT,MPOINT,G
      INTEGER NGPOINT,NMPOINT,NGBLOCK,NMBLOCK,GMCI,M,I
      INTEGER NMGROUP,NGGROUP,MTRIGE,MTRIGP,LOWMFILT,HIGHMFILT
      INTEGER MTRIGL,GFILTER,MFILTER,GEVELT,MEVELT,GEVLOC,MEVLOC
      INTEGER GEVTAB(12,19),MEVTAB(12,38),MMCI(108)
      DOUBLE PRECISION DGTAPTME,DMTAPTME,GGPTME,MGPTME,WINDOW
      DOUBLE PRECISION GEVTME,MEVTME,MGP2ADD
      REAL GBLOCK(8192),MBLOCK(8192),GGROUP(612),MGROUP(1220)
      DATA NGBLOCK,NMBLOCK,GMCI /3*0/
      DATA MMCI /108*0/
      COMMON GEVTAB,MEVTAB,GGROUP,MGROUP
      WINDOW=0.0045D0


      WRITE(*,*) 'OPENING RESULTS FILES'
      OPEN (51, FILE='/home/orion/casd/disk3/CC30TS.DAT'
     +,STATUS='UNKNOWN')
```

```
      WRITE(51,*) 'RESULTS OF GLASGOW/MUNICH COINCIDENCE SEARCH'
C      OPEN (52, FILE='/home/orion/casd/disk3/CCSN.DAT'
C     +,STATUS='UNKNOWN')
      WRITE(52,*) 'SN EVENTS IN MUNICH DATA'
      OPEN (53, FILE='/home/orion/casd/disk3/CC30CB1.DAT'
     +,STATUS='UNKNOWN')
      OPEN (54, FILE='/home/orion/casd/disk3/CC30CB2.DAT'
     +,STATUS='UNKNOWN')
      OPEN (55, FILE='/home/orion/casd/disk3/CC30CB3.DAT'
     +,STATUS='UNKNOWN')
      OPEN (56, FILE='/home/orion/casd/disk3/CC30CB4.DAT'
     +,STATUS='UNKNOWN')
      OPEN (57, FILE='/home/orion/casd/disk3/CC30CB5.DAT'
     +,STATUS='UNKNOWN')
      OPEN (58, FILE='/home/orion/casd/disk3/CC30CB6.DAT'
     +,STATUS='UNKNOWN')
      OPEN (59, FILE='/home/orion/casd/disk3/CC30CB7.DAT'
     +,STATUS='UNKNOWN')
      OPEN (60, FILE='/home/orion/casd/disk3/CC30CB8.DAT'
     +,STATUS='UNKNOWN')
      OPEN (61, FILE='/home/orion/casd/disk3/CC30CB9.DAT'
     +,STATUS='UNKNOWN')
      OPEN (62, FILE='/home/orion/casd/disk3/CC30CB10.DAT'
     +,STATUS='UNKNOWN')

      GTAPNO=0
      MTAPNO=0
      NGBLOCK=0
      NMBLOCK=0
C     THESE ARE SET TO 1 BY THE TESTL ALGORITHMS


      CALL TESTLGB(GBLOCK,GTAPNO,DGTAPTME,NGBLOCK)
      WRITE(*,*) 'LOADING 1ST GBLOCK'
      WRITE(*,*) 'DGTAPTME=',DGTAPTME
      WRITE(*,*) 'LOADING NGBLOCK',NGBLOCK
      WRITE(*,*) 'SUBROUTINE TESTLGB FINISHED'
      WRITE(*,*) 'LOADING 1ST MBLOCK'
      CALL TESTLMB(MBLOCK,MTAPNO,DMTAPTME,NMBLOCK)
      WRITE(*,*) 'DMTAPTME=',DMTAPTME
      WRITE(*,*) 'LOADING NMBLOCK',NMBLOCK
      WRITE(*,*) 'SUBROUTINE TESTLMB FINISHED'
           WRITE(*,*) 'WRITING GGROUP TO ARRAY'
      DO 20 I=1,612
         GGROUP(I)=GBLOCK(I)
20    CONTINUE
      WRITE (*,*) 'CALLING GTABLE'
      CALL GTABLE
C     CREATE TABLE OF LISTS OF EVENTS FOR THE VARIOUS FILTERS, AND
C     SEND THEM ACROSS AS A COMMON BLOCK
C     WRITE(*,*) 'GEVTAB ARRAY IS',GEVTAB

      WRITE(*,*) 'WRITING MGROUPS TO ARRAY'
      DO 30 I=1,1220
```

```
            MGROUP(I)=MBLOCK(I)
   30     CONTINUE
          NGGROUP=1
          NMGROUP=2
          WRITE(*,*) 'CALLING MTABLE'
          CALL MTABLE
C         WRITE(*,*) 'MEVTAB ARRAY IS',MEVTAB
          WRITE(*,*) 'GLASGOW TAPETIME IS',DGTAPTME
          WRITE(*,*) 'MUNICH TAPETIME IS',DMTAPTME
          GGPTME=DGTAPTME
          MGPTME=DMTAPTME
c         NOTE THAT MGPTME IS SET TO DMTATTME FOR THE FIRST DUAL GROUP

******WRITTEN GLASGOW BLOCK1 GROUP1 TO ARRAY GGROUP, MUNICH BLOCK1 GROUPS1
******AND 2 TO ARRAY MGROUP


***********************************************************************
   50     WRITE(*,*) '#####################################'
          WRITE(*,*) 'SETUP COMPLETED; BEGINNING ANALYSIS FOR'
          WRITE(*,*) 'TS EVENTS (FILTER NO. 1)'
          WRITE(*,*) 'AND FOR NGGROUP,NMGROUP',NGGROUP,NMGROUP
          GFILTER=1
          MFILTER=1
C         GO TO START OF MGROUP, WHETHER OR NOT IT IS NEW, BECAUSE IT
C         WOULD BE MESSY TO TRY TO START WHERE WE LEFT WITH FOR PREVIOUS GGROUP
          MEV=1
          IF (MEVTAB(MFILTER,MEV).EQ.0) GO TO 220
          MTRIGE=1
          MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
C          WRITE(*,*) 'MEVELT=',MEVELT
C         MEVELT IS SHORTHAND FOR THIS EXPRESSION ON RHS. DECLARE WHENEVER
C         MFILTER OR MEV CHANGES.
          NMPOINT=MGROUP(MEVELT+3)-MGROUP(MEVELT+1)+1
C         THE NUMBER OF DATA POINTS IN THIS MUNICH EVENT
          MPOINT=1

          WRITE(*,*) 'STARTING ANALYSIS OF G AND M DATA'
          GFILTER=1
          MFILTER=1
          GEV=0
   60     GEV=GEV+1
             GMCI=0
C             WRITE(*,*) 'GFILTER,GEV=',GFILTER,GEV
C             WRITE(*,*) 'GEVTAB(GFILTER,GEV)=',GEVTAB(GFILTER,GEV)
             IF ( (GEVTAB(GFILTER,GEV).EQ.0).OR.(GEV.GT.19) ) THEN
C                WRITE(*,*) 'STARTING NEXT FILTER'
                GO TO 220
             ENDIF
             GEVELT=GEVLOC(GEVTAB(GFILTER,GEV))
C             WRITE(*,*) 'GEVLOC(GEVTAB(GFILTER,GEV)=',GEVELT
             MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
C             WRITE(*,*) 'GEVELT AND MEVELT ARE', GEVELT,MEVELT
C            REPLACE LENGTHY ARRAY ADDRESSES WITH GEVELT,MEVELT
```

```
C            START NEXT FILTER
 NGPOINT=GGROUP(GEVELT+3)-GGROUP(GEVELT+1)+1
C     THE NUMBER OF DATA POINTS IN THIS GLASGOW EVENT
C             WRITE(*,*) 'GEVELT+3=',GEVELT+3,' MEVELT=',MEVELT
C  WRITE(*,*) 'NUMBER OF DATA POINTS IN THIS GEVENT',NGPOINT
         GPOINT=1
80  IF ( (GEVTAB(GFILTER,GEV).EQ.0).OR.(GEV.GT.19) ) THEN
C             WRITE(*,*) 'STARTING NEXT FILTER'
            GO TO 220
         ENDIF
C     WRITE(*,*) 'BEGINNING TIME COMPARISON FOR GEV,MEV',GEV,MEV
C         WRITE(*,*) 'FOR TS FILTER'
C         WRITE(*,*) 'GPOINT AND MPOINT OF THESE EVS:', GPOINT,MPOINT
C         WRITE(*,*) 'GMCI=',GMCI
C         WRITE(*,*) 'MMCI FOR THIS MEV =',MMCI(MEVTAB(MFILTER,MEV))
C  WRITE(*,*) 'GGPTME=',GGPTME,'MGPTME=',MGPTME
C  WRITE(*,*) 'GEV,MEV AT TIMES'
C  WRITE(*,*) (GGROUP(GEVELT+1)+GPOINT-1)/20000.0
100  GEVTME=DBLE((GGROUP(GEVELT+1)+GPOINT-1)/20000.0)
C  WRITE(*,*) '(DOUBLE PRECISION VALUE',GEVTME+GGPTME,')'
c        ADD 3.0 TO MEVTME IF MEV IN 2ND GROUP OF DUAL GROUP
         IF (MEVTAB(MFILTER,MEV).GT.54) THEN
            MGP2ADD=3.0D0
         ELSE
            MGP2ADD=0.0D0
         ENDIF
C         WRITE(*,*) (MGROUP(MEVELT+1)+MPOINT-1)/10000.0
 MEVTME=DBLE((MGROUP(MEVELT+1)+MPOINT-1)/10000.0)+MGP2ADD
C  WRITE(*,*) '(DOUBLE PRECISION VALUE',MEVTME+MGPTME,')'
c  WRITE(*,*) 'WINDOW=',WINDOW
c  WRITE(*,*) GEVTME+GGPTME-WINDOW,MEVTME+MGPTME
c  WRITE(*,*) GEVTME+GGPTME+WINDOW,MEVTME+MGPTME
 IF ((GEVTME+GGPTME-WINDOW) .LT. (MEVTME+MGPTME)) THEN
    IF ((GEVTME+GGPTME+WINDOW) .GT. (MEVTME+MGPTME)) THEN

C        COINCIDENCE!
       WRITE(*,*) 'COINCIDENCE'
C              WRITE GGROUP(GEV) AND MGROUP(MEV) TO FILE, WITH THE
C              GROUP HEADERS, THRESHOLD EVENT MINI HEADERS AND DELAY.
C              ALSO STORE MTRIGGER, THE FIRST MUNICH EVENT TO COINCIDE
C              WITH THIS GLASGOW EVENT, TO RETURN TO IT WHEN FINISHED
C              WITH THE GLASGOW EVENT.
       GMCI=GMCI+1
C
               MMCI(MEVTAB(MFILTER,MEV))=MMCI(MEVTAB(MFILTER,MEV))+1
C        WRITE(*,*) 'GMCI=',GMCI
C               WRITE(*,*) 'MMCI=',MMCI(MEVTAB(MFILTER,MEV))
C        GMCI IS GLASGOW MULTIPLE COINCIDENCE INDEX, I.E. THE
C              NUMBER OF COINCIDENCES INVOLVING EACH GPOINT
C        MMCI(MEVTAB(MFILTER,MEV)) IS MUNICH MULTIPLE COINCIDENCE
C              INDEX FOR EVENT MEV; MEVTAB RUNS FROM 1 TO 108.
               IF (GMCI .EQ. 1) THEN
C   WRITE(*,*) 'MEV',MEV,'IS TRIGGER FOR GEV',GEV
```

```
C   WRITE(*,*) 'AT GPOINT=',GPOINT,'MPOINT=',MPOINT
                  MTRIGE=MEV
C                  WRITE(*,*) 'MEV,MTRIGE=',MEV,MTRIGE
  MTRIGL=NMPOINT
  MTRIGP=MPOINT
       ENDIF

               WRITE(51,1900) GEVTME+GGPTME
C        EXPERIMENTAL TIME OF GLASGOW POINT IN QUESTION
       WRITE(51,*) GGROUP(1),NGBLOCK,NGGROUP,GGPTME
               WRITE(51,*) (GGROUP(G),G=8,12)
               WRITE(51,*) (GGROUP(G),G=14,17)
               WRITE(51,*) (GGROUP(GEVELT+G),G=1,3),
     +                      GGROUP(GEVELT+1)+GPOINT-1
       WRITE(51,*) GGROUP(GEVELT),GMCI
       WRITE(51,*) (GGROUP(GEVELT+G),G=4,5)
       WRITE(51,*) (GGROUP(GEVELT+G),G=6,9)
               WRITE(51,1900) MEVTME+MGPTME
       WRITE(51,*) MGROUP(1),NMBLOCK,NMGROUP,MGPTME
               WRITE(51,*) (MGROUP(M),M=8,12)
               WRITE(51,*) (MGROUP(M),M=214,217)
               WRITE(51,*) (MGROUP(MEVELT+M),M=1,3),
     +                      MGROUP(MEVELT+1)+MPOINT-1
               WRITE(51,*) MGROUP(MEVELT),MMCI(MEVTAB(MFILTER,MEV))
       WRITE(51,*) (MGROUP(MEVELT+M),M=4,5)
       WRITE(51,*) (MGROUP(MEVELT+M),M=6,9)
       WRITE(51,1900)
     +   (DNINT((GEVTME+GGPTME-MEVTME-MGPTME)*1D5))/1D5
C        WRITE(*,*) (GEVTME+GGPTME-MEVTME-MGPTME)*1D5
C        WRITE(*,*) (DNINT((GEVTME+GGPTME-MEVTME-MGPTME)*1D5))/1D5
C        WRITE(*,*) 'IS THE DELAY'
       WRITE(51,*) '#'

               MPOINT=MPOINT+1
               IF (MPOINT .GT. NMPOINT) THEN
                  MEV=MEV+1
                  IF ( (MEVTAB(MFILTER,MEV).EQ.0).OR.(MEV.EQ.39) ) THEN
c                    THAT WAS THE LAST MEV OF THIS DUAL MGROUP FOR THIS FILTER
                     IF (GMCI.EQ.0) GO TO 220
C                     I.E. THERE IS NO TRIGGER EVENT TO RETURN TO
C                     WRITE(*,*) 'NO MORE COINC POSS'
C                     WRITE(*,*) 'FOR THIS GPOINT: TAKE NEXT'
C                     WRITE(*,*) 'GPOINT AND RETURN TO TRIGGER'
                     GPOINT=GPOINT+1
                     GMCI=0
C              GO BACK TO MUNICH TRIGGER POINT
C                     WRITE(*,*) 'RETURNING TO TRIGGER EVENT',MTRIGE
                     MEV=MTRIGE
             NMPOINT=MTRIGL
             MPOINT=MTRIGP
                     MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
                     IF (GPOINT .GT. NGPOINT) THEN
        GO TO 60
                     ELSE
```

```
                          GO TO 100
                       ENDIF
                    ENDIF
                    MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
                    NMPOINT=MGROUP(MEVELT+3)-MGROUP(MEVELT+1)+1
C   THE NUMBER OF DATA POINTS IN THIS NEW MUNICH EVENT
                    MPOINT=1
                 ENDIF
                 GO TO 100

              ELSE
150              GPOINT=GPOINT+1
C                IF THE LAST POINTS WERE COINCIDENTAL WE WANT TO RETURN
C                TO THE TRIGGER EVENT; BUT IF THEY WERE NOT WE CAN CONTINUE
C                WITH THE CURRENT MPOINT.
                 IF (GMCI.EQ.0) GO TO 170
          GMCI=0
C           GO BACK TO MUNICH TRIGGER POINT
C                 WRITE(*,*) 'RETURNING TO TRIGGER EVENT',MTRIGE
                 MEV=MTRIGE
          NMPOINT=MTRIGL
          MPOINT=MTRIGP
                 MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
170              IF (GPOINT .GT. NGPOINT) THEN
  GO TO 60
                 ELSE
                    GO TO 100
                 ENDIF
              ENDIF
           ELSE
              MPOINT=MPOINT+1
              IF (MPOINT .GT. NMPOINT) THEN
                 MEV=MEV+1
                 IF ( (MEVTAB(MFILTER,MEV).EQ.0).OR.(MEV.EQ.39) ) THEN
c                   THAT WAS THE LAST MEV OF THIS DUAL MGROUP FOR THIS FILTER
                    IF (GMCI.EQ.0) GO TO 220
C                    I.E. IF THERE IS NO TRIGGER EVENT TO RETURN TO
C                    WRITE(*,*) 'NO MORE COINCIDENCES POSS FOR THIS GPOINT'
C                    WRITE(*,*) 'SO TAKE NEXT GPOINT AND RETURN TO TRIGGER'
                    GPOINT=GPOINT+1
                    GMCI=0
C             GO BACK TO MUNICH TRIGGER POINT
C                    WRITE(*,*) 'RETURNING TO TRIGGER EVENT',MTRIGE
                    MEV=MTRIGE
            NMPOINT=MTRIGL
            MPOINT=MTRIGP
                    MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
                    IF (GPOINT .GT. NGPOINT) THEN
                       GO TO 60
                    ELSE
                       GO TO 100
                    ENDIF
                 ENDIF
                 MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
```

```
                  NMPOINT=MGROUP(MEVELT+3)-MGROUP(MEVELT+1)+1
C                 THE NUMBER OF DATA POINTS IN THIS NEW MUNICH EVENT
                  MPOINT=1
               ENDIF
               GO TO 100
            ENDIF

200      GO TO 60

C        FINISHED WITH THIS GFILTER FOR THIS GGROUP

**********************************************************************

C        SEARCH FOR VERY UNUSUAL SN TYPES IN MUNICH DATA
C        NOW MFILTER=2 FOR THIS MGROUP. OLD MGROUP WILL STILL BE APPROPRIATE.
C        CANNOT PERFORM COINCIDENCE ANALYSIS WITH NO SIMILAR GLASGOW DATA.

220      WRITE(*,*) 'SKIPPING ANALYSIS FOR FILTER 2'
         GO TO 410



**********************************************************************

C        ANALYSIS OF GLASGOW FILTER MATCHES BY GFILTERS 3-12 ONE BY ONE
410      WRITE(*,*) 'STARTING COALESCING BINARY FILTER MATCHES'
         DO 700 GFILTER=3,12
C         WRITE(*,*) '***NEW GFILTER: NUMBER IS',GFILTER
C        CHOOSE THE CORRECT MFILTERS FOR EACH GFILTER
         IF (GFILTER.EQ.3) THEN
            LOWMFILT=3
            HIGHMFILT=4
         ENDIF
         IF ((GFILTER.GE.4).AND.(GFILTER.LE.11)) THEN
            LOWMFILT=GFILTER-1
            HIGHMFILT=GFILTER+1
         ENDIF
         IF (GFILTER.EQ.12) THEN
            LOWMFILT=11
            HIGHMFILT=12
         ENDIF
         DO 650 MFILTER=LOWMFILT,HIGHMFILT
C         WRITE(*,*) '***NEW MFILTER: NUMBER IS',MFILTER

         MEV=1
         GMCI=0
         MTRIGE=1

C         WRITE(*,*) 'MEVTAB(MFILTER,MEV)=',MEVTAB(MFILTER,MEV)
         IF (MEVTAB(MFILTER,MEV).EQ.0) GO TO 650
         MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
         NMPOINT=MGROUP(MEVELT+7)-MGROUP(MEVELT+6)+1
C         WRITE(*,*) 'NMPOINT=',NMPOINT
C        THE NUMBER OF DATA POINTS IN THIS MUNICH EVENT
```

```
        MPOINT=1

  450   GEV=0
C        WRITE(*,*) 'EXECUTED 450'
  460   GEV=GEV+1
C           WRITE(*,*) 'NEW GEV'
C           WRITE(*,*) 'GEV=',GEV
C           WRITE(*,*) 'GFILTER=',GFILTER
C           WRITE(*,*) 'GEVTAB(GFILTER,GEV)=', GEVTAB(GFILTER,GEV)
           IF ( (GEVTAB(GFILTER,GEV).EQ.0).OR.(GEV.GT.16) )THEN
C              WRITE(*,*) 'GEVTAB(GFILTER,GEV) IS 0 JUST AFTER 460 SO'
C              WRITE(*,*) 'STARTING NEXT MFILTER'
              GO TO 650
           ENDIF

           GEVELT=GEVLOC(GEVTAB(GFILTER,GEV))
           MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
C          REPLACE LENGTHY ARRAY ADDRESSES WITH GEVELT,MEVELT
 GMCI=0
C           WRITE(*,*) 'RESETTING GMCI BECAUSE NEW GEV'
 NGPOINT=GGROUP(GEVELT+7)-GGROUP(GEVELT+6)+1
C     THE NUMBER OF DATA POINTS IN THIS GLASGOW EVENT
C  WRITE(*,*) 'NUMBER OF DATA POINTS IN THIS G EVENT',NGPOINT
           GPOINT=1

  480  IF (GEVTAB(GFILTER,GEV).EQ.0) THEN
C              WRITE(*,*) 'STARTING NEXT MFILTER'
              GO TO 650
           ENDIF

C  WRITE(*,*) 'BEGINNING TIME COMPARISON FOR GEV,MEV',GEV,MEV
C          WRITE(*,*) 'FOR COALESCING BINARY GFILTER',GFILTER
C  WRITE(*,*) 'GPOINT AND MPOINT OF THESE EVS:', GPOINT,MPOINT
C  WRITE(*,*) 'GGPTME=',GGPTME,'MGPTME=',MGPTME
C  WRITE(*,*) 'GEV,MEV AT TIMES'
C  WRITE(*,*) (GGROUP(GEVELT+6)+GPOINT-1)/20000.0
  500  GEVTME=DBLE((GGROUP(GEVELT+6)+GPOINT-1)/20000.0)
C  WRITE(*,*) '(DOUBLE PRECISION VALUE',GEVTME+GGPTME,')'
c          ADD 3.0 TO MEVTME IF MEV IN 2ND GROUP OF DUAL GROUP
           IF (MEVTAB(MFILTER,MEV).GT.54) THEN
              MGP2ADD=3.0D0
           ELSE
              MGP2ADD=0.0D0
           ENDIF
c  WRITE(*,*) (MGROUP(MEVELT+6)+MPOINT-1)/10000.0
 MEVTME=DBLE((MGROUP(MEVELT+6)+MPOINT-1)/10000.0)+MGP2ADD
c         WRITE(*,*) '(DOUBLE PRECISION VALUE',MEVTME+MGPTME,')'

C          WRITE MASSPAR AND PHASE TO SCREEN
C           WRITE(*,*) 'GMASSPAR=',GGROUP(GEVELT+4)
C           WRITE(*,*) 'GPHASE=',GGROUP(GEVELT+5)
C           WRITE(*,*) 'MMASSPAR=',MGROUP(MEVELT+4)
```

```
C        WRITE(*,*) 'MPHASE=',MGROUP(MEVELT+5)

c  WRITE(*,*) 'WINDOW=',WINDOW
c  WRITE(*,*) GEVTME+GGPTME-WINDOW,MEVTME+MGPTME
c  WRITE(*,*) GEVTME+GGPTME+WINDOW,MEVTME+MGPTME
 IF ((GEVTME+GGPTME-WINDOW) .LT. (MEVTME+MGPTME)) THEN
    IF ((GEVTME+GGPTME+WINDOW) .GT. (MEVTME+MGPTME)) THEN

C         COINCIDENCE!
        WRITE(*,*) 'COINCIDENCE'
C               WRITE GGROUP(GEV) AND MGROUP(MEV) TO FILE, WITH THE
C               GROUP HEADERS, THRESHOLD EVENT MINI HEADERS AND DELAY.
C               ALSO STORE MTRIGGER, THE FIRST MUNICH EVENT TO COINCIDE
C               WITH THIS GLASGOW EVENT, TO RETURN TO IT WHEN FINISHED
C               WITH THE GLASGOW EVENT.
        GMCI=GMCI+1
                MMCI(MEVTAB(MFILTER,MEV))=MMCI(MEVTAB(MFILTER,MEV))+1
C         WRITE(*,*) 'GMCI=',GMCI
C                WRITE(*,*) 'MMCI FOR THIS MEV=',MMCI(MEVTAB(MFILTER,MEV))
C         GMCI IS GLASGOW MULTIPLE COINCIDENCE FLAG FOR EACH GPOINT
        IF (GMCI .EQ. 1) THEN
C   WRITE(*,*) 'MEV',MEV,'IS TRIGGER FOR GEV',GEV
C   WRITE(*,*) 'AT GPOINT=',GPOINT,'MPOINT=',MPOINT
                   MTRIGE=MEV
   MTRIGL=NMPOINT
   MTRIGP=MPOINT
        ENDIF
C                WRITE(*,*) (DBLE(GGROUP(GEVELT+6)+GPOINT-1)/2D4)+GGPTME
        WRITE(50+GFILTER,1900) GEVTME+GGPTME
C         EXPERIMENTAL TIME OF GLASGOW POINT IN QUESTION
        WRITE(50+GFILTER,*) GGROUP(1),NGBLOCK,NGGROUP,GGPTME
                WRITE(50+GFILTER,*) (GGROUP(G),G=8,12)
                WRITE(50+GFILTER,*) (GGROUP(G),G=14,17)
                WRITE(50+GFILTER,*) GGROUP(GEVELT+6),GGROUP(GEVELT+3),
     +             GGROUP(GEVELT+7),GGROUP(GEVELT+6)+GPOINT-1
        WRITE(50+GFILTER,*) GGROUP(GEVELT),GMCI
        WRITE(50+GFILTER,*) (GGROUP(GEVELT+G),G=1,2)
                WRITE(50+GFILTER,*) GGROUP(GEVELT+4),GGROUP(GEVELT+5)
        WRITE(50+GFILTER,*) (GGROUP(GEVELT+G),G=8,11)
        WRITE(50+GFILTER,1900) MEVTME+MGPTME
        WRITE(50+GFILTER,*) MGROUP(1),NMBLOCK,NMGROUP,MGPTME
                WRITE(50+GFILTER,*) (MGROUP(M),M=8,12)
                WRITE(50+GFILTER,*) (MGROUP(M),M=14,17)
                WRITE(50+GFILTER,*) MGROUP(MEVELT+6),MGROUP(MEVELT+3),
     +             MGROUP(MEVELT+7),MGROUP(MEVELT+6)+MPOINT-1
                WRITE(50+GFILTER,*) MGROUP(MEVELT),
     +             MMCI(MEVTAB(MFILTER,MEV))
        WRITE(50+GFILTER,*) (MGROUP(MEVELT+M),M=1,2)
                IF (MGROUP(MEVELT+4).GT.2.0) THEN
                WRITE(50+GFILTER,*) MGROUP(MEVELT+4)-4.0,
     +             MGROUP(MEVELT+5)
                ELSE
                WRITE (50+GFILTER,*) MGROUP(MEVELT+4),
     +             MGROUP(MEVELT+5)
```

```
C              CORRECTING FOR ERROR IN RECORDED MMASSPAR
               ENDIF
C!             MEVELT(2) IS h INFORMATION: USE IT HERE.
      WRITE(50+GFILTER,*) (MGROUP(MEVELT+M),M=8,11)
      WRITE(50+GFILTER,1900)
     +  (DNINT((GEVTME+GGPTME-MEVTME-MGPTME)*1D5))/1D5
C       WRITE(*,*) (GEVTME+GGPTME-MEVTME-MGPTME)*1D5
C       WRITE(*,*) (DNINT((GEVTME+GGPTME-MEVTME-MGPTME)*1D5))/1D5
C              ALL THE NINTS AND 1D5'S ARE TO REMOVE APPARENT NUM. ERRORS
C       WRITE(*,*) 'IS THE DELAY'
               WRITE(50+GFILTER,*) '#'

               MPOINT=MPOINT+1
               IF (MPOINT .GT. NMPOINT) THEN
C                  WRITE(*,*) 'EXECUTING 520'
520               MEV=MEV+1
                  IF ( (MEVTAB(MFILTER,MEV).EQ.0).OR.(MEV.EQ.33) ) THEN
C                    THAT WAS THE LAST MEV OF THIS DUAL GROUP FOR THIS FILTER
                     IF (GMCI.EQ.0) GO TO 650
C                     WRITE(*,*) 'NO MORE COINC POSS FOR THIS GPOINT:'
C                     WRITE(*,*) 'TAKE NEXT GPOINT AND RETURN TO TRIGGER'
                     GPOINT=GPOINT+1
                     GMCI=0
C                    GO BACK TO MUNICH TRIGGER POINT
C                     WRITE(*,*) 'RETURNING TO TRIGGER EVENT',MTRIGE
                     MEV=MTRIGE
                     NMPOINT=MTRIGL
                     MPOINT=MTRIGP
                     MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
                     IF (GPOINT .GT. NGPOINT) THEN
      GO TO 460
                     ELSE
                       GO TO 500
                     ENDIF
                  ENDIF
                  MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
                  NMPOINT=MGROUP(MEVELT+7)-MGROUP(MEVELT+6)+1
C                  WRITE(*,*) 'NMPOINT=',NMPOINT
C   THE NUMBER OF DATA POINTS IN THIS NEW MUNICH EVENT
                  MPOINT=1
               ENDIF
               GO TO 500

            ELSE
550            GPOINT=GPOINT+1
      GMCI=0
C       GO BACK TO MUNICH TRIGGER POINT

C       WRITE(*,*) 'RETURNING TO TRIGGER EVENT',MTRIGE
               MEV=MTRIGE
               MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
      NMPOINT=MTRIGL
      MPOINT=MTRIGP
               IF (GPOINT .GT. NGPOINT) THEN
```

```
 GO TO 460
              ELSE
                 GO TO 500
              ENDIF
           ENDIF
        ELSE
           MPOINT=MPOINT+1
           IF (MPOINT .GT. NMPOINT) THEN
C           WRITE(*,*) 'EXECUTING 570'
570        MEV=MEV+1
              IF ( (MEVTAB(MFILTER,MEV).EQ.0).OR.(MEV.EQ.33) ) THEN
C                THAT WAS THE LAST MEV OF THIS DUAL GROUP FOR THIS FILTER
                 IF (GMCI.EQ.0) GO TO 650
C                 WRITE(*,*) 'NO MORE COINC POSS FOR THIS GPOINT:'
C                 WRITE(*,*) 'TAKE NEXT GPOINT AND RETURN TO TRIGGER'
                 GPOINT=GPOINT+1
                 GMCI=0
C                GO BACK TO MUNICH TRIGGER POINT
C                 WRITE(*,*) 'RETURNING TO TRIGGER EVENT',MTRIGE
                 MEV=MTRIGE
                 NMPOINT=MTRIGL
                 MPOINT=MTRIGP
                 MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
                 IF (GPOINT .GT. NGPOINT) THEN
                    GO TO 460
                 ELSE
                    GO TO 500
                 ENDIF
              ENDIF
              MEVELT=MEVLOC(MEVTAB(MFILTER,MEV))
              NMPOINT=MGROUP(MEVELT+7)-MGROUP(MEVELT+6)+1
C              WRITE(*,*) 'NMPOIT=',NMPOINT
C              WRITE(*,*) 'MEVELT=',MEVELT
C             THE NUMBER OF DATA POINTS IN THIS NEW MUNICH EVENT
              MPOINT=1
           ENDIF
           GO TO 500
 ENDIF

600   GO TO 460

650   CONTINUE
C     NEXT MFILTER
700   CONTINUE
C     NEXT GFILTER

C     USED ALL GEVs AND GFILTERS FOR THIS GGROUP
C     GET NEW GGROUP AND NEW GBLOCK, MGROUP AND MBLOCK IF NECESSARY
**********************************************************************


C     NEW GGROUP; NEW GBLOCK IF NECESSARY

1000  NGGROUP=NGGROUP+1
```

```
      IF (NGGROUP .GT. 13) THEN
         WRITE(*,*) 'LOADING NEW GLASGOW BLOCK'
         NGGROUP=1
         CALL TESTLGB(GBLOCK,GTAPNO,DGTAPTME,NGBLOCK)
         WRITE(*,*) 'DGTAPTME=',DGTAPTME
      ENDIF
C     NEW GGROUP ANYWAY
      WRITE(*,*) 'CREATING ARRAY FOR GGROUP',NGGROUP
1010  DO 1020 I=1,612
         GGROUP(I)=GBLOCK(612*(NGGROUP-1)+I)
1020  CONTINUE
      IF (GGROUP(1).EQ.0) THEN
C        GLASGOW TAPE ENDS MIDWAY THROUGH A RESULT BLOCK
C         WRITE(*,*) 'EXECUTING IF OPTION AT LINE 1020'
         NGGROUP=1
         CALL TESTLGB(GBLOCK,GTAPNO,DGTAPTME,NGBLOCK)
         WRITE(*,*) 'DGTAPTME=',DGTAPTME
         GO TO 1010
      ENDIF
      CALL GTABLE
      GGPTME=DGTAPTME+
     +   DBLE((NGGROUP-1)*1.3104)+((NGBLOCK-1)*13*1.3104)
      WRITE(*,*) 'GGPTME=',GGPTME


      IF ( ((MGPTME+2*3.0)-(GGPTME+1.3104)) .LT. WINDOW) THEN

C        NEW MUNICH DATA
C        NEW MBLOCK IF NECESSARY
1100     NMGROUP=NMGROUP+1
         IF (NMGROUP .GT. 13) THEN
C     LOAD FRESH MBLOCK.
    NMGROUP=1
            WRITE(*,*) 'LOADING FRESH MBLOCK'
            CALL TESTLMB(MBLOCK,MTAPNO,DMTAPTME,NMBLOCK)
         WRITE(*,*) 'DMTAPTME=',DMTAPTME
         ENDIF

C        NEW MGROUP ANYWAY
1110     DO 1120 I=1,610
            MGROUP(I)=MGROUP(I+610)
1120     CONTINUE
1130     DO 1140 I=1,610
            MGROUP(I+610)=MBLOCK((610*(NMGROUP-1))+I)
1140     CONTINUE
         IF (MGROUP(611).EQ.0) THEN
C     MUNICH TAPE ENDS MIDWAY THROUGH A RESULT BLOCK
C     WRITE(*,*) 'EXECUTING IF OPTION AT LINE 1140'
    NMGROUP=1
    CALL TESTLMB(MBLOCK,MTAPNO,DMTAPTME,NMBLOCK)
    WRITE(*,*) 'DMTAPTME=',DMTAPTME
    GO TO 1130
         ENDIF
C        RESHUFFLE MMCI
```

```
            DO 1150 I=1,54
               MMCI(I)=MMCI(I+54)
 1150       CONTINUE
            DO 1160 I=1,54
               MMCI(I+54)=0
 1160       CONTINUE
C            WRITE(*,*) 'MMCI=',MMCI


            MGPTME=DMTAPTME+
     +        DBLE((NMGROUP-2)*3.0)+((NMBLOCK-1)*13*3.0)
c           THIS IS THE TIME OF THE FIRST GROUP OF THE PAIR, WHEREAS
C           NMGROUP REFERS TO THE SECOND OF THE PAIR
            WRITE(*,*) 'DMTAPTME=',DMTAPTME,'NMBLOCK=',NMBLOCK
            WRITE(*,*) 'NMGROUP=',NMGROUP,'MEV=',MEV
            WRITE(*,*) 'MGPTME=',MGPTME
            CALL MTABLE

         ENDIF

         GO TO 50

C        THE PROGRAM WILL NOT REACH THIS STAGE. RATHER, IT WILL FAIL AT
C        ONE OF THE LOAD SUBROUTINES, WHEN THERE IS NOTHING TO LOAD

 1900     FORMAT (F15.6)

         END




         SUBROUTINE TESTLGB(GBLOCK,GTAPNO,DGTAPTME,NGBLOCK)

         REAL GBLOCK(8192),EGBLOCK(8192)
         REAL GTAPINT,GTAPSEC
         DOUBLE PRECISION DGTAPTME,DGTAPSEC
         INTEGER GTAPNO,NGBLOCK,GBLREC
         DATA GBLREC/0/

         OPEN(80,FILE='/home/orion/dxn/chris/glasgow_30to40',
     +    STATUS='OLD',ACCESS='DIRECT',RECL=32768,
     +    FORM='UNFORMATTED')
         NGBLOCK=NGBLOCK+1
         GBLREC=GBLREC+1


         READ(80,REC=GBLREC) EGBLOCK
C        THIS STUFF ABOUT EGBLOCK IS FOR SIMULATION OF THE CONVERSION
C        FROM CHARACTER FORMAT DATA TO REAL NUMBERS, REQUIRED LATER IF
C        DATA READ FROM EXABYTES

         WRITE(*,*) 'READING GBLOCK NUMBER', NGBLOCK
         DO 100 I=1,8192
```

```
 GBLOCK(I)=EGBLOCK(I)
100   CONTINUE

      IF (GBLOCK(1).NE.GTAPNO) THEN
 GTAPNO=GBLOCK(1)
 GTAPINT=((GBLOCK(4)-2.0)*24*60*60)
     +  +((GBLOCK(5)-15.0)*60*60) +(GBLOCK(6)*60)
 GTAPSEC=GBLOCK(7)
C  WRITE(*,*) 'GTAPSEC=',GTAPSEC
 DGTAPSEC=( (DBLE(NINT(GTAPSEC*1D5)))/1D5 )+3.2D-3
C  WRITE(*,*) 'DGTAPSEC=',DGTAPSEC
 DGTAPTME=DBLE(GTAPINT)+DGTAPSEC
C        DGTAPTME IS THE TIME IN SECONDS SINCE 2d14h0m0s UT
 NGBLOCK=1
C        RESET NGBLOCK SO THAT GGPTME IS CALCULATED WRT TAPTME
      ENDIF


      RETURN
      END
```

```
      SUBROUTINE TESTLMB(MBLOCK,MTAPNO,DMTAPTME,NMBLOCK)

      REAL MBLOCK(8192),EMBLOCK(8192)
      REAL MTAPINT,MTAPSEC
      INTEGER MTAPNO,NMBLOCK,MBLREC
      DOUBLE PRECISION DMTAPTME,DMTAPSEC
      DATA MBLREC/16/
      OPEN (90,FILE='/home/orion/dxn/chris/munich_30to40',
     +    STATUS='OLD',ACCESS='DIRECT',RECL=32768,
     +    FORM='UNFORMATTED')
      NMBLOCK=NMBLOCK+1
      MBLREC=MBLREC+1

      READ (90,REC=MBLREC) EMBLOCK


C     THIS STUFF ABOUT EMBLOCK IS FOR SIMULATION OF THE CONVERSION
C     FROM CHARACTER FORMAT DATA TO REAL NUMBERS, REQUIRED LATER IF
C     DATA READ FROM EXABYTES
      DO 100 I=1,8192
 MBLOCK(I)=EMBLOCK(I)
100   CONTINUE

      IF(MBLOCK(1).NE.MTAPNO) THEN
 MTAPNO=MBLOCK(1)
 MTAPINT=((MBLOCK(4)-2.0)*24*60*60)
     +    +((MBLOCK(5)-15.0)*60*60)+((MBLOCK(6)*60))
     +    -3600
c        THIS -3600 IS AN IMMEDIATE NORMALISATION TO UNIVERSAL TIME
```

```
663  MTAPSEC=MBLOCK(7)
664  DMTAPSEC=(DBLE(NINT(MTAPSEC*1D5)))/1D5
665  DMTAPTME=DBLE(MTAPINT)+DMTAPSEC
666 c        DMTAPTME IS THE TIME IN SECONDS SINCE 2d14h0m0s UT ie 15h LOCAL
667  WRITE(*,*) 'DMTAPTME=',DMTAPTME
668  NMBLOCK=1
669 c     RESET TO 1 AT EVERY NEW TAPE, SO AS TO GET MGPTME RIGHT
670 C     AT LINE 1160
671       ENDIF

672
673       RETURN
674       END



675       FUNCTION GEVLOC(GEVENT)

676 C     FUNCTION SUBPROGRAM TO LOCATE THE 1ST ELEMENT OF THE NEXT EVENT
677 C     IN THE ARRAY GGROUP. THE EVENTS ''GEVENT'' ARE THE GROUPS
678 C     OF TEN OR TWELVE REAL NUMBERS CORRESPONDING TO TIME SERIES EVENTS
679 C     OR FILTER MATCHES, IN THE ORDER IN WHICH THEY OCCUR IN GGROUP.
680 C     THE RETURNED FUNCTION VALUE EVLOC(GEVENT) IS THE POSITION OF THE
681 C     START OF THE EVENT GEVENT IN ARRAY GGROUP.

682       INTEGER GEVLOC,GEVENT

683 ******SET POINTERS TO THE EVENTS IN THE ORDER IN WHICH THEY OCCUR IN
684 ******ARRAY GGROUP
685
686 C      WRITE(*,*) 'DOING A GEVLOC FOR GEVENT',GEVENT

687       IF (GEVENT.GT.19) THEN
688          IF (GEVENT.GT.35) THEN
689             IF (GEVENT.GT.51) THEN
690                WRITE(*,*) 'ERROR: ARRAY GGROUP EXHAUSTED'
691                STOP
692             ELSE
693                GEVLOC=412+1+1+(GEVENT-35-1)*12
694             ENDIF
695          ELSE
696             GEVLOC=212+1+1+(GEVENT-19-1)*12
697          ENDIF
698       ELSE
699 C        WRITE(*,*) 'CALCULATING GEVLOC AT EXPECTED PLACE'
700         GEVLOC=(GEVENT-1)*10+19+1
701 C        WRITE(*,*) 'GEVLOC AS CALCULATED BY GEVLOC SUBTN. IS',GEVLOC
702       ENDIF

703       END
```

```
      FUNCTION MEVLOC(MEVENT)

C     FUNCTION SUBPROGRAM TO LOCATE THE 1ST ELEMENT OF THE NEXT EVENT
C     IN THE ARRAY MGROUP. THE EVENTS ''IEVENT'' ARE THE GROUPS
C     OF TEN OR TWELVE REAL NUMBERS CORRESPONDING TO TIME SERIES EVENTS
C     OR FILTER MATCHES, IN THE ORDER IN WHICH THEY OCCUR IN GGROUP.
C     THE RETURNED FUNCTION VALUE EVLOC(IEVENT) IS THE POSITION OF THE
C     START OF THE EVENT IN ARRAY MGROUP.

      INTEGER MEVLOC,MEVENT

******SET POINTERS TO THE EVENTS IN THE ORDER IN WHICH THEY OCCUR IN
******ARRAY MGROUP

C      WRITE(*,*) 'DOING A MEVLOC FOR MEVENT', MEVENT
      IF (MEVENT.GT.19) THEN
         IF (MEVENT.GT.38) THEN
            IF (MEVENT.GT.54) THEN
               IF (MEVENT.GT.73) THEN
                  IF (MEVENT.GT.92) THEN
                     IF (MEVENT.GT.108) THEN
                        WRITE(*,*) 'ERROR: ARRAY MGROUP EXHAUSTED'
                        STOP
                     ELSE
                        MEVLOC=1022+1+1+(MEVENT-92-1)*12
                     ENDIF
                  ELSE
                     MEVLOC=822+7+1+(MEVENT-73-1)*10
                  ENDIF
               ELSE
                  MEVLOC=622+7+1+(MEVENT-54-1)*10
               ENDIF
            ELSE
               MEVLOC=412+1+1+(MEVENT-38-1)*12
            ENDIF
         ELSE
            MEVLOC=212+7+1+(MEVENT-19-1)*10
         ENDIF
      ELSE
        MEVLOC=12+7+1+(MEVENT-1)*10
      ENDIF

      END


      SUBROUTINE GTABLE

C     THIS SUBROUTINE IS TO PUT ALL THE EVENTS IN THE ARRAY GGROUP
C     INTO A TABLE CLASSIFYING THEM BY THE FILTER WITH WHICH THEY
C     MATCH, REGARDING TIME SERIES EVENTS AS MATCHES WITH A TRIVIAL
C     FILTER, NO. 1. THUS THERE ARE 10 CB FILTERS, ONE TS
C     FILTER AND ONE 2.5 kHz SN FILTER. THE MAXIMUM NUMBER OF EVENTS
C     FOR ONE FILTER IS 19 (16 FOR CB OR SN).
```

```
C     THE TABLE GEVTAB IS PASSED TO
C     THE MAIN PROGRAM BY A COMMON STATEMENT.
C     FUNCTION GEVLOC IS CALLED TO LOCATE THE GEVENT'TH EVENT.

      INTEGER GEVTAB(12,19),MEVTAB(12,38)
C     COLUMN 1 IS TIME SERIES; COL 2 IS SN (BLANK FOR THE GLASGOW
C     DATA); COLS 3-12 ARE CB.
      INTEGER GEVENT,GFILTER,GLISTELT,I,J,GEVLOC,GFFF
      REAL GGROUP(612),MGROUP(1220)
      REAL GMASSPAR
      COMMON GEVTAB,MEVTAB,GGROUP,MGROUP

C      WRITE(*,*) 'GGROUP IS'
C      WRITE(*,*) GGROUP
      DO 60 I=1,12
         DO 50 J=1,19
            GEVTAB(I,J)=0
50       CONTINUE
60    CONTINUE
      WRITE(*,*) 'ARRAY GEVTAB SET UP AND ZEROED'

C     GGROUP(GEVLOC(GEVENT)) IS THE ELEMENT IN GGROUP CORRESPONDING
C     TO THE START OF THE GEVENT'TH EVENT (IN ORDER OF OCCURRENCE IN GGROUP).
      DO 200 GEVENT=1,51
C         WRITE(*,*) 'GEVENT=',GEVENT
         GFFF=GEVLOC(GEVENT)
C         WRITE(*,*) 'GEVLOC(GEVENT)=',GFFF
C        FORBIDDEN:  WRITE(*,*) 'GEVLOC(GEVENT)=',GEVLOC(GEVENT)
C         WRITE(*,*) 'GGROUP(GEVLOC(GEVENT))=',GGROUP(GFFF)
         IF (GGROUP(GFFF) .EQ. 0) THEN
C              WRITE(*,*) 'THIS IS A NON-EVENT'
            GO TO 200
         ENDIF

C        CHOOSE APPROPRIATE FILTER LIST IN TABLE FOR THIS EVENT
         IF (GEVENT.LE.19) THEN
C           TIME SERIES EVENT
C            WRITE(*,*) 'THIS IS A TS EVENT'
C           FIND THE FIRST ZERO-VALUED ELEMENT IN COLUMN 1 AND PUT IN GEVENT
            GLISTELT=1
100         IF (GEVTAB(1,GLISTELT).EQ.0) THEN
               GEVTAB(1,GLISTELT)=GEVENT
C               WRITE(*,*) 'GEVTAB(1,GLISTELT) NOW CONTAINS', GEVENT
               GO TO 200
            ELSE
               GLISTELT=GLISTELT+1
               IF (GLISTELT.GT.19) THEN
                  WRITE(*,*) 'ERROR: MORE THAN 19 TS GEVENTS
     +            FOUND IN GGROUP'
                  STOP
               ENDIF
               GO TO 100
            ENDIF
         ENDIF
```

```

            IF ((GEVENT.GE.20) .AND. (GEVENT.LE.35)) THEN
C                WRITE(*,*) 'THIS IS A 1.4M CB EVENT'
C               COALESCING BINARY EVENT IN SUBGROUP 2, I.E. HAVING MASS
C               PARAMETER NEAR 1.4 SOLAR MASSES.
C               FIND APPROPRIATE COLUMN IN WHICH TO PUT GEVENT

                GMASSPAR=GGROUP(GEVLOC(GEVENT)+4)
C                WRITE(*,*) 'MASS PARAMETER IS',GMASSPAR

C               GMASSPAR MUST BE 1.3761+MULTIPLE OF 0.00695 (x 1 to 5)
                GFILTER=NINT(((GMASSPAR-1.3761)/0.00695)+2.0)
C               GFILTERS STORED IN INCREASING MASS FROM 1.38305 TO 1.4178
C               AS GFILTER RUNS FROM 3 TO 7

                GLISTELT=1
  140           IF (GEVTAB(GFILTER,GLISTELT).EQ.0) THEN
                   GEVTAB(GFILTER,GLISTELT)=GEVENT
                   GO TO 200
                ELSE
                   GLISTELT=GLISTELT+1
                   IF (GLISTELT.GT.16) THEN
                      WRITE(*,*) 'ERROR: MORE THAN 16  1.4M CB
     +                   GEVENTS FOUND IN GGROUP'
                      STOP
                   ENDIF
                   GO TO 140
                ENDIF
             ENDIF

             IF (GEVENT.GE.36 .AND. GEVENT.LE.51) THEN
C                WRITE(*,*) 'THIS IS AN INTEGRAL MASS CB EVENT'
C               COALESCING BINARY EVENT IN SUBGROUP 2, I.E. HAVING
C               INTEGER MASS PARAMETER IN INTERVAL [2,6]
C               FIND APPROPRIATE COLUMN IN WHICH TO PUT GEVENT
                GMASSPAR=GGROUP(GEVLOC(GEVENT)+4)
C                WRITE(*,*) 'MASS PARAMETER IS',GMASSPAR
                GFILTER=NINT(GMASSPAR+6.0)
C                WRITE(*,*) 'GMASSPAR+6.0 IS'
C                WRITE(*,*) GMASSPAR+6.0
C                WRITE(*,*) 'NINT(GMASSPAR+6.0) IS'
C                WRITE(*,*) NINT(GMASSPAR+6.0)
C                WRITE(*,*) 'GFILTER=', GFILTER

                GLISTELT=1
  120           IF (GEVTAB(GFILTER,GLISTELT).EQ.0) THEN
                   GEVTAB(GFILTER,GLISTELT)=GEVENT
                   GO TO 200
                ELSE
                   GLISTELT=GLISTELT+1
                   IF (GLISTELT.GT.16) THEN
                      WRITE(*,*) 'ERROR: MORE THAN 16 INTEGER
     +                     MASS CB GEVENTS FOUND IN GGROUP'
                      STOP
```

```
                   ENDIF
                   GO TO 120
                ENDIF
             ENDIF

    200   CONTINUE
          GEVENT=51

    C     ALL GEVENTS IN THIS GGROUP DISTRIBUTED IN CORRECT COLUMNS OF GEVTAB
    C      WRITE(*,*) 'GTABLE VERSION OF GEVTAB IS',GEVTAB
          RETURN
          END


          SUBROUTINE MTABLE

    C     THIS SUBROUTINE IS TO PUT ALL THE EVENTS IN THE ARRAY MGROUP
    C     INTO A TABLE CLASSIFYING THEM BY THE FILTER WITH WHICH THEY
    C     MATCH, REGARDING TIME SERIES EVENTS AS MATCHES WITH A TRIVIAL
    C     FILTER, NO. 1. THUS THERE ARE 10 CB FILTERS, ONE TS
    C     FILTER AND ONE 2.5kHz SN FILTER. THE MAXIMUM NUMBER OF EVENTS
    C     FOR ONE FILTER IS 38 (32 FOR CB OR SN).
    C     THE TABLE MEVTAB IS PASSED TO
    C     THE MAIN PROGRAM BY A COMMON STATEMENT.
    C     FUNCTION MEVLOC IS CALLED TO LOCATE THE MEVENT'TH EVENT.

          INTEGER GEVTAB(12,19),MEVTAB(12,38)
    C     COLUMN 1 IS TIME SERIES; COL 2 IS SN; COLS 3-12 ARE CB.
          INTEGER MEVENT,MFILTER,MLISTELT,I,J,MEVLOC,MFFF
          REAL GGROUP(612),MGROUP(1220)
          REAL MMASSPAR
          COMMON GEVTAB,MEVTAB,GGROUP,MGROUP

          DO 60 I=1,12
             DO 50 J=1,38
                MEVTAB(I,J)=0
    50       CONTINUE
    60    CONTINUE
          WRITE(*,*) 'ARRAY MEVTAB SET UP AND ZEROED'

    C     MGROUP(MEVLOC(MEVENT)) IS THE ELEMENT IN MGROUP CORRESPONDING
    C     TO THE START OF THE MEVENT'TH EVENT (IN ORDER OF OCCURRENCE IN MGROUP).
          DO 200 MEVENT=1,108
    C         WRITE(*,*) 'MEVENT=',MEVENT
             MFFF=MEVLOC(MEVENT)
    C         WRITE(*,*) 'MEVLOC(MEVENT)=',MFFF
    C         WRITE(*,*) 'MGROUP(MEVLOC(MEVENT))=',MGROUP(MFFF)
             IF (MGROUP(MFFF) .EQ. 0) THEN
    C            WRITE(*,*) 'THIS IS A NON-EVENT'
                GO TO 200
             ENDIF

    C        CHOOSE APPROPRIATE FILTER LIST IN TABLE FOR THIS EVENT
             IF ((MEVENT.LE.19) .OR. ((MEVENT.GE.55).AND.(MEVENT.LE.73)))
```

```
     +   THEN
C            WRITE(*,*) 'THIS IS A SN EVENT'
C           CANDIDATE SUPERNOVA EVENT
C           FIND THE FIRST ZERO-VALUED ELEMENT IN COLUMN 2 AND PUT IN MEVENT
            MLISTELT=1
100         IF (MEVTAB(2,MLISTELT).EQ.0) THEN
               MEVTAB(2,MLISTELT)=MEVENT
               GO TO 200
            ELSE
               MLISTELT=MLISTELT+1
               IF (MLISTELT.GT.38) THEN
                  WRITE(*,*) 'ERROR: MORE THAN 38 SUPERNOVA CANDIDATES
     +            FOUND IN MGROUP'
                  STOP
               ENDIF
               GO TO 100
            ENDIF
         ENDIF

         IF (((MEVENT.GE.20).AND.(MEVENT.LE.38)).OR.
     +   ((MEVENT.GE.74).AND.(MEVENT.LE.92))) THEN
C            WRITE(*,*) 'THIS IS A TS EVENT'
C           TIME SERIES EVENT
C           FIND THE FIRST ZERO-VALUED ELEMENT IN COLUMN 1 AND PUT IN MEVENT
            MLISTELT=1
120         IF (MEVTAB(1,MLISTELT).EQ.0) THEN
               MEVTAB(1,MLISTELT)=MEVENT
               GO TO 200
            ELSE
               MLISTELT=MLISTELT+1
               IF (MLISTELT.GT.38) THEN
                  WRITE(*,*) 'ERROR: MORE THAN 38 TIME SERIES EVENTS
     +            FOUND IN MGROUP'
                  STOP
               ENDIF
               GO TO 120
            ENDIF
         ENDIF

         IF (((MEVENT.GE.39).AND.(MEVENT.LE.54)).OR.
     +   ((MEVENT.GE.93).AND.(MEVENT.LE.108))) THEN
C            WRITE(*,*) 'THIS IS A CB EVENT'
C           COALESCING BINARY EVENT
C           FIND APPROPRIATE COLUMN IN WHICH TO PUT MEVENT
            MMASSPAR=MGROUP(MEVLOC(MEVENT)+4)
            IF (MMASSPAR.GT.2.0) MMASSPAR=MMASSPAR-4.0
C!          CORRECTING FOR ERROR IN ORIGINAL RECORDING OF MUNICH MASSPAR
C            WRITE(*,*) 'MASS PARAMETER IS',MMASSPAR
            IF (MMASSPAR .LT. 1.5) THEN
C              MMASSPAR MUST BE 1.3761+MULTIPLE OF 0.00695
               MFILTER=NINT(((MMASSPAR-1.3761)/0.00695)+2.0)
C              MFILTERS STORED IN INCREASING MASS.
C              MASS PARAMETER ONLY (NOT PHASE) IS THE
C              ONLY INDICATOR AS TO WHETHER A CANDIDATE COINCIDENCE HAS BEEN
```

```
C               FOUND. SIMILARLY FOR GTABLE
             ELSE
C               MMASSPAR MUST BE INTEGER VALUE IN [2,6]
                MFILTER=NINT(MMASSPAR+6.0)
             ENDIF

             MLISTELT=1
C            WRITE(*,*) 'MFILTER=',MFILTER,'MLISTELT=',MLISTELT
140          IF (MEVTAB(MFILTER,MLISTELT).EQ.0) THEN
                MEVTAB(MFILTER,MLISTELT)=MEVENT
                GO TO 200
             ELSE
                MLISTELT=MLISTELT+1
                IF (MLISTELT.GT.32) THEN
                   WRITE(*,*) 'ERROR: MORE THAN 32 CB MEVENTS FOUND IN
     +             MGROUP'
                   STOP
                ENDIF
                GO TO 140
             ENDIF
          ENDIF

200   CONTINUE
      MEVENT=108

C     ALL MEVENTS IN THIS MGROUP DISTRIBUTED IN CORRECT COLUMNS OF MEVTAB
C      WRITE(*,*) 'MTABLE VERSION OF MEVTAB IS', MEVTAB
      RETURN
      END
```

Appendix B

Comments on the 100 Hour Experiment

I include here a list of my comments on the 100 hour experiment, so that any future experimenters or analysis groups can perhaps learn from my experience. The comments are entirely my own. I am keeping this list informal, in the interests of keeping it complete. Most of the comments and recommendations are common sense, and many might even consider them obvious. Undeterred, I list them here.

Although I have split the list into "data analysis" and "experimental" sub-lists, much of this will be of interest to both data analysers and experimenters. Furthermore, although some of the comments seem to pertain exclusively to the 100 Hour Experiment, I am sure that most of this will be of interest to any future experimental and analysis teams.

Particular to data analysis:

1. **Volume of data** The large volume of data used was difficult to work with. At times it was even prohibitive (e.g. it being difficult and time-consuming to return to the original data to check the sign of the gravitational wave threshold-crossing events.) This was partly due to the intrinsically huge amount of disk space which the original datasets would have occupied ($\sim 100 \times 3600 \times 20000 \sim 7.2 \times 10^9$ real numbers in the Glasgow Secondary Error Point stream alone), but also partly due to the storage formats used. Although an Exabyte tape can store about three hours of output from the Glasgow detector, including all the housekeeping streams, access to the data, as well as the time taken to read it, was slow. Initially, it was also found that these tape drives were prone

to stoppages, jams, resettings and even breakdowns (particularly for Watkins). By the time of my research, these problems had become much less frequent, at least during the reading and copying of the results tapes which Watkins and Nicholson had written.

The time taken in the changing of tapes during the experiment was also a problem, causing gaps in the data. The changing of tapes at the analysis stage was inconvenient, but not as important. I also found some repetitions of data on the tapes, and it is not clear at which stage this occurred.

I think it is inevitable that, when the large detectors are observing there will be some sort of heirarchy of data storage. One will require fast access for (a) data taken recently; (b) a dataset reduced in some way (e.g. by the setting of a threshold or by the operation of Fourier transforms, etc.); and (c) data of ongoing interest. For data archived for the long term, the priority would be high density of data storage, perhaps at the expense of speed of access. At this moment, I would not recommend Exabytes to be used for the former task; a faster access format would be more desirable: perhaps optical disks or optical paper. For the longer term storage, Exabytes, DAT and particularly the new Terabyte tape drive look promising. Of course, this situation is likely to change.

In general, I found the event lists could be stored on a smaller number of tapes (potentially one or two tapes for the event list from each detector) than the original data tapes (about 30 Exabyte tapes for Glasgow, about 100 1/2-inch tapes for Munich); while the coincidence list could easily be stored on the hard disk of a desktop computer. This resulted in progressively easier and faster and more convenient access at each stage of the analysis.

2. **Unpacking of data** In the running of the original data analysis routines, to look for events, the 100 hours of data took many times longer than 100 hours to analyse. Certainly, the unpacking of the data (reading and demultiplexing the data on the original tape) was an analysis bottleneck. This was particularly true for Watkins (1991), though undoubtedly contributing to this were the large number of data streams recorded by Glasgow, and the problems Watkins experienced with the Exabyte tape drives. I believe the analysis stage could be considerably speeded up if the unpacking were made quicker and easier. Could this be done without reducing the amout of data recorded? Perhaps the structuring of the data could be made easier to read at the

analysis stage without making it excessively difficult to write at the detector.

3. **Glasgow Calibration** The calibration of the Glasgow data proved to be quite troublesome at the analysis stage, with various stages and factors to be considered. Experimenters should bear in mind the analysis overheads of calibration when designing the data output system. Can the calibration could be done in real time at the detector?

4. **Glasgow gravitational wave output** While Watkins (Watkins 1991) simply used the Secondary Error Point signal as the gravitational wave signal output of the Glasgow detector, it is known that it is preferable to combine the Secondary Error Point and Secondary Feedback signals, which are more sensitive in different frequency ranges (Robertson 1990). This should be taken care of in future analyses of the 100 Hour dataset.

   As for future experiments:

5. **Format of storage of the events lists** Event lists should have as similar structure as possible. By this I mean that there should be the same number of elements in each block, each block should cover the same length of time in the experiment, and so on. The coincidence analysis of the 100 Hour Experiment, especially the writing of the coincidence program, would have been easier if the two event lists had been stored with the same structure (see e.g. Watkins 1991).

6. **Information relating to each event** In compiling the event lists, *all* the information commensurate to each event in the original tapes should be retained. This applies, firstly, to the housekeeping data, where I have noted in Chapter 7 that Watkins dispensed with some housekeeping data (Watkins 1991) which may have turned out to be useful. It also applies where the two individual analysis programs have recorded only the peak values of signal-to-noise for an event which may be several datapoints long. I believe that the signal-to-noise should be recorded for all datapoints in an event, as: (a) event morphology is important for physics, and (b) it could later be used to construct a duration veto (see Chapter 7).

   Of course, this would somewhat increase the amount of data retained at the individual analysis stage, but hopefully this would be outweighed by improvements in storage/retrieval and processing power.

7. **Oscillatory bursts** In the individual analysis programs, it would be helpful if long, oscillatory bursts in the output were recorded as one event, rather than several discrete threshold crossing events. Some thought is required as to how a program could differentiate between, on the one hand, oscillatory events and, on the other hand, clusters of consecutive and genuinely independent threshold crossers. Perhaps a warning flag should be attached to suspicious-looking event clusters?

8. **Sign of event** The individual analysis should record the sign of the signal-to-noise of events, not just the absolute value. Again, this is useful both from the physical point of view, and from the vetoing point of view, since it could single-handedly reduce the number of spurious coincidences by up to 50%.

9. **Sensitivity veto** The use of a sensitivity veto, or *sigma veto*, would considerably improve the limits set by any future test experiments. If used discriminately, it should not result in the loss of too much data. (See Chapter 7 in this thesis).

10. **Application of vetoes** As I have said in Chapter 7, I believe that housekeeping data would be much more effective for vetoing if the vetoes were applied *on line*, and based on sudden changes or on a running distribution recorded relatively recently (e.g. over ten minutes, rather than based on a distribution taken over one hundred hours). To do this, there would need to be a preliminary analysis system working in real time with the operation of the detector, perhaps at the detector site. If this is impossible, one would need to record and bin potentially every datum output by the detector.

11. **General analysis system** I believe that the threshold – event list – coincidence comparison method adopted in the analysis so far has been successful and, given some refinements, should be fairly straightforward. If this were the main method adopted for analysis of data from the first generation of long interferometers, I think it would be desirable to have two tiers in the analysis. The first tier would be a preliminary analysis particular to each detector, with thresholds and filters, performed on-site. The reduced dataset could then be transmitted, via satellite or optical fibre, to a central specialist data analysis site; and there, a comparison analysis, similar to the one performed in this thesis, could be made. This is not a new idea (see e.g. Schutz 1988a; Corbett 1988; Schutz 1991), but my experience with this analysis has shown at least that such a two-tier method works in principle.

Particular to the experimenters:

12. **Gaps in data** 10% of the Glasgow data was useless due to losses of lock, while about the same fraction of the Munich data was lost due to tape changes. However, I believe both of these situations will improve. The Glasgow prototype is not currently designed with automatic relock (recovery of resonance in the cavities), which may change. I hope also that for future working observatories, the tape changes will either be practically instantaneous, or that there will be two tape drives attached, so that at least one is working at all times.

13. **Garching housekeeping data** In future experiments, the Garching laboratory should record more housekeeping data, particularly microphone information and an AC current glitch detector. These would be useful anyway, but would be particularly helpful for any future problems with laser water-cooler switching. The group at Garching should also, if possible, set the water-cooling threshold higher; thus making it switch less often, as pointed out by Rüdiger (1992).

    Again, the increase in the volume of the data thus incurred should be outweighed by future improvements in data storage and processing power.

14. **Dwangies** I think some attention should be directed by the experimenters to understanding and preventing the "dwangies": short, high amplitude oscillatory bursts, seen in the output of the three main prototype detectors. In the time series, these look remarkably like gravitational wave sources are expected to look, and thus will raise the false alarm rate.

    This also pertains to the analysis groups: is there some housekeeping information, perhaps one which we are already recording, which could be used to remove dwangies? Either they must be eliminated, or a reliable veto must be found. The analysis and experimental groups should discuss this.

And finally:

15. **Communication between groups** I believe that the communication between the experimental groups and the analysis group in Cardiff could have been better. This is, of course, the fault of no-one. Although it has improved in the last year, I believe it could be improved yet again, with perhaps more reciprocal visits between groups. I have

found through experience that much time can be saved during analysis, by consulting members of the appropriate groups. My attitude is not to be shy; collaborations exist so that people can collaborate, even at the risk of wasting someone else's time. In my consultations with, and faxes and e-mail messages to, various members of the experimental groups, I have received nothing but courteous, friendly, and prompt assistance. This has markedly speeded up my coincidence analysis, and has been gratefully received.

References

Abramowitz, Milton & Stegun, Irene A., 1972. Handbook of mathematical functions, Dover Pub. Inc., New York, 8th edition.

Aglietta, M., Badino, G., Bologna, G., Castagnoli, C., Castellina, A., Dadykin, V. L., Fulgione, W., Galeotti, P., Kalchukov, F. F., Kortchaguin, B., Kortchaguin, P.V., Malguin, A. S., Ryassny, V. G., Ryazhskaya, O.G., Saavedra, O., Talochkin, V. P., Trinchero, G., Vernetto, S., Zatsepin, G. T. & Yakushev, V. F., 1987a, On the event observed in the Mont Blanc Underground Neutrino Observatory during the occurrence of Supernova SN1987a. *Europhysics Letters*, **3**, 1315.

Aglietta, M., Badino, G., Bologna, G., Castagnoli, C., Castellina, A., Dadykin, V. L., Fulgione, W., Galeotti, P., Kalchukov, F. F., Kortchaguin, B., Kortchaguin, P.V., Malguin, A.S., Ryassny, V.G., Ryazhskaya, O.G., Saavedra, O., Talochkin, V.P., Trinchero, G., Vernetto, S., Zatsepin, G.T. & Yakushev, V.F., 1987b, Comments on the two events observed in neutrino detectors during the supernova 1987a outburst. *Europhysics Letters*, **3**, 1321.

Aglietta, M., Badino, G., Bologna, G., Castagnoli, C., Castellina, A., Fulgione, W., Galeotti, P., Saavedra, O., Trinchero, G., Vernetto, S., Amaldi, E., Cosmelli, C., Frasca, S., Pallottino, G.V., Pizzella, G., Rapagnani, P., Ricci, F., Bassan, M., Coccia, E., Modena, I., Bonifazi, P., Castellano, M. G., Dadykin, V.L., Malguin, A. S., Ryassny, V.G., Ryazhskaya, O. G., Yakushev, V. F., Zatsepin, G. T., Gretz, D., Weber, J. & Wilmot, G., 1989, Analysis of the data recorded by the Mont Blanc neutrino detector and by the Maryland and Rome gravitational-wave detectors during SN1987A. *Il Nuovo Cimento*, **12C**, 75.

Aglietta, M., Castellina, A., Fulgione, W., Trinchero, G., Vernetto, S., Astone, P., Badino, G., Bologna, G., Bassan, M., Coccia, E., Modena, I., Bonifazi, P., Castellano, M.G., Visco, M., Castagnoli, C., Galeotti, P., Saavedra, O., Cosmelli, C., Frasca, S., Pallottino, G.V., Pizzella, G., Rapagnani, P., Ricci, F., Majorana, E., Gretz, D., Weber, J. & Wilmot, G., 1991a, Correlation between the Maryland and Rome gravitational wave detectors and the Mont Blanc, Kamioka and IMB particle detectors during SN1987A. *Il Nuovo Cimento*, **106B**, 1257.

Aglietta, M., Castellina, A., Fulgione, W., Trinchero, G., Vernetto, S., Castagnoli, C., Galeotti, P., Saavedra, O., Amaldi, E., Frasca, S., Pallottino, G. V., Pizzella, G., Rapagnani, P., Ricci, F., Astone, P., Cosmelli, C., Bassan, M., Coccia, E., Modena, I., Bonifazi, P., Castellano, M.G., Visco, M., Badino, G., Bologna, G., Dadykin, V.L., Khalchukov, F.F., Korolkova, I.V., Kortchaguin, P.V., Kudryatzev, V.A., Malguin, A.S., Ryassny, V.G., Ryazhskaya, O.G., Yakushev, V.F., Zatsepin, G.T., Gretz, D., Weber, J. & Wilmot, G., 1991b, Coincidences among the data recorded by the Baksan, Kamioka and Mont Blanc underground neutrino detectors, and by the Maryland and Rome gravitational-wave detectors during Supernova 1987 A. *Il Nuovo Cimento*, **14C**, 171.

Alexeyev, E.N., Alexeyeva, L.N., Krivosheina, I. V. & Volchenko, V.I., 1987, Possible detection of a neutrino signal on 23 February 1987 at the Baksan underground scintillation telescope of the Institute-of-Nuclear-Research. *JETP Letters*, **45**, 589.

Amaldi, E., Bonifazi, P., Castellano, M.G., Coccia, E., Cosmelli, C., Frasca, S., Gabellieri, M., Modena, I., Pallottino, G.V., Pizzella, G., Rapagnani, P., Ricci, F. & Vannaroni, G., 1987, Data recorded by the Rome room temperature gravitational wave antenna, during the Supernova SN 1987a in the Large Magellanic Cloud. *Europhysics Letters*, **3**, 1325.

Amaldi, E., Bonifazi, P., Frasca, S., Gabellieri, M., Gretz, D., Pallottino, G.V., Pizzella, G., Weber, J. & Wilmot, G., 1988. Analysis of the data recorded by the Maryland and Rome room temperature gravitational wave antennas in the period of the SN 1987A, In: *Fourth George Mason University Workshop in Astrophysics: SN1987A in the LMC*, 453, eds Kafatos, M. & Michalitsianos, A., Cambridge University Press, New York.

Amaldi, E., Bassan, M., Bonifazi, P., Castellano, M.G., Coccia, E., Cosmelli, C., Frasca, S., Gretz, D., Modena, I., Pallottino, G.V., Pizzella, G., Rapagnani, P., Ricci, F., Weber,

J. & Wilmot, G, 1989. Coincidences among the Maryland and Rome gravitational wave detector data and the Mont Blanc and Kamioka neutrino detector data in the period of SN1987A, In: *Proceedings of the Fourteenth Texas Symposium on Relativistic Astrophysics*, 571, ed. Fenyves, E.J.

Arp, H.A., 1987. Quasars, Redshifts and Controversies, Interstellar Media, Berkely, California.

Astone, P., Bassan, M., Bonifazi, P., Carelli, P., Castellano, M.G., Cavallari, G., Coccia, E., Cosmelli, C., Fafone, V., Frasca, S., Modena, E. Majorana I., Pallottino, G.V., Pizzella, G., Rapagnani, P., Ricci, F. & Visco, M., 1993, Long term operation of the Rome "Explorer"cryogenic gravitational wave detector. *Physical Review D*, **47**, 362.

Astone, P., Lobo, J.A. & Schutz, B.F., 1993. Coincidence experiments between interferometric and resonant bar detectors of gravitational waves, Preprint, University of Wales College of Cardiff; submitted to Physical Review D.

Astone, P., Private communication. 1992. Subject: Rome–Stanford bar coincidence analysis.

Bionta, R. M., Blewitt, G., Bratton, C. B., Casper, D., Ciocio, A., Claus, R., Cortez, B., Crouch, M., Dye, S. T., Errede, S., Foster, G. W., Gajewski, W., Ganezer, K. S., Goldhaber, M., Haines, T. J., Jones, T. W., Kielczewska, D., Kropp, W. R., Learned, J. G., LoSecco, J. M., Mattews, J., Miller, R., Mudan, M. S., Park, H. S., Price, L. R., Reines, F., Shultz, J., Seidel, S., Shumard, E., Sinclair, D., Sobel, H. W., Stone, J. L., Sulak, L. R., Svoboda, R., Thornton, G., van der Velde, J. C. & Wuest, C., 1987, Observations of a neutrino burst in coincidence with Supernova 1987A in the Large Magellanic Cloud. *Physical Review Letters*, **58**, 1494.

Bulmer, M.G., 1979. Principles of Statistics, Dover, New York.

Chudakov, A.E., 1989. On the Correlation between Mont Blanc and Baksan Underground Detectors in February 1987, In: *Fourteenth Texas Symposium on Relativistic Astrophysics*, 577, ed. Fenyves, E.J., Kluwer Academic Publishers, Dordrecht.

Corbett, I.F., 1988. Round-table on data exchange, In: *Gravitational wave data analysis: proceedings of the NATO advanced research workshop held at Dyffryn House, St. Nicholas, Cardiff, Wales, 6-9 July 1987*, p. 327, ed. Schutz, B.F., Kluwer Academic Publishers, Dordrecht.

Cutler, Curt, Apostolatos, Theocharis A., Bildsten, Lars, Finn, Lee Samuel, Flanagan, Eanna E., Kennefick, Daniel, Markovic, Dragoljubov M., Ori, Amos, Poisson, Eric, Sussman, Gerald Jay & Thorne, Kip S., 1993, The last three minutes: Issues in gravitational-wave measurements of coalescing compact binaries. *Physical Review Letters*, **70**, 2984.

Davis, M.H.A, 1988. A Review of the Statistical Theory of Signal Detection, In: *Gravitational wave data analysis: proceedings of the NATO advanced research workshop held at Dyffryn House, St. Nicholas, Cardiff, Wales, 6-9 July 1987*, p. 73, ed. Schutz, B.F., Kluwer Academic Publishers, Dordrecht.

Dickson, C.A. & Schutz, Bernard F., 1992. Reassessment of the reported correlations between gravitational waves and neutrinos associated with SN1987A, Preprint, University of Wales College of Cardiff. Submitted to Physical Review D.

Einstein, A. In: *Preuss. Akad. Wiss. Berlin*, 688, Sitzungsberichte der Physikalischmathematischen Klasse, 1916.

Faller, J.E., Hollander, W.J., Nelson, P.G. & McHugh, M.P., 1990, Gyroscope-weighing experiment with a null result. *Physical Review Letters*, **64**, 825.

Gavryuseva, E.A., Gavryusev, V.G. & Rosljakov, A.G., 1990. Statistical analysis of neutrino capture rate in ^{37}Cl experiment, Preprint, Institute for Nuclear Research of the Academy of Sciences of the USSR.

Grishchuk, L.P., 1992, Quantum mechanics of a solid-state bar gravitational wave antenna. *Physical Review D*, **45**, 2601.

Guthrie, B.N.G. & Napier, W.M., 1991, Evidence for redshift periodicity in nearby field galaxies. *Mon.Not.R.A.S*, **253**, 533.

Hayasaka, H. & Takeuchi, S., 1989, Anomalous weight reduction on a gyroscope's right rotations around the vertical axis on the earth. *Physical Review Letters*, **63**, 2701.

Hirata, K., Kajita, T., Koshiba, M., Nakahata, M., Oyama, Y., Sato, N., Suzuki, A., Takita, M., Totsuka, Y., Kifune, T., Suda, T., Takahashi, K., Tanimori, T., Miyano, K., Yamada, M., Beier, E.W., Feldscher, L. R., Kim, S. B., Mann, A. K., Newcomer, F. M., Berg, R. Van, Zhang, W. & Cortez, B. G., 1987, Observation of a neutrino burst from the supernova SN1987A. *Physical Review Letters*, **58**, 1490.

Hough, J., Meers, B.J., Newton, G.P., Robertson, N.A., Ward, H., Schutz, B.F., Drever, R.W.P, Mason, R., Pollard, C., Tolcher, R., Bellenger, D.W., Bennett, J.R.J, Corbett, I.F. & Percival, M.D., 1986. Technical Report, A British Long Baseline Gravitational Wave Observatory, University of Glasgow.

Hough, J., Meers, B.J., Newton, G.P., Robertson, N.A., Ward, H., Leuchs, G., Niebauer, T.M., Rüdiger, A., Schilling, R., Schnupp, L., Walther, H., Winkler, W., Schutz, B.F., Ehlers, J., Kafka, P., Schäfer, G., Hamilton, M.W., Schütz, I, Welling, H., Bennett, J.R.J., Corbett, I.F, Edwards, B.W.H., Greenhalgh, R.J.S & Kose, V., 1989. Technical Report, Proposal for a Joint German–British Interferometric Gravitational Wave detector, Max-Planck-Institut für Quantenoptik, Garching, Germany.

Hough, J., Kerr, G.A., Mackenzie, N.L., Mangan, J.B., Meers, B.J., Newton, G.P., Robertson, D.I., Robertson, N., Ward, H., Danzmann, K., Niebauer, T.M., Rüdiger, A., Schilling, R., Schnupp, L., Winkler, W., Dickson, C.A., Jones, G.S., Nicholson, D., Schutz, B.F., Shuttleworth, J. & Watkins, W.J., 1993. Results of the first coincident observations by two gravitational wave interferometers, Preprint, University of Wales College of Cardiff.

Newton, G, 1993. Workshop Summary: Gravitational wave experiments; summary compiled by W.O. Hamilton and Ho Jung Paik, In: *Genaral Relativity and Gravitation 1992: Proceedings of the Thirteenth International Conference on General Relativity and Gravitation, 1992*, p. 407, eds Gleiser, R.J., Kozameh, C.N. & Moreschi, O.M., Institute of Physics Publishing, Bristol.

Nicholson, David, Schutz, Bernard F., Shuttleworth, Justin R. & Watkins, W. John, 1992. Progress in Gravitational Wave Detection, In: *Proceedings of the Sixth Marcel Grossmann Meeting on General Relativity; Kyoto, Japan, 23-29 June 1991*, p. 163, eds Sato, Humitaka & Nakamura, Takashi, World Scientific, Singapore.

Nicholson, D., Private communication. 1991. Subject: general.

Nicholson, D., Private communication. 1993. Subject: general, 1.25 kHz low-pass filter.

Niebauer, T.M., Rüdiger, A., Schilling, R., Schnupp, L., Winkler, W. & Danzmann, K., 1993, Pulsar search using data-compression with the Garching gravitational-wave detector. *Physical Review D*, **47**, 3106.

Phinney, E.S., 1991, The rate of neutron star binary mergers in the Universe: minimal predictions for gravity wave detectors. *Astrophysical Journal*, **380**, L17.

Pizzella, G. & Pallottino, G., Private communication. 1991.

Pizzella, G., 1988, Coincidence Techniques for Gravitational Wave Experiments. *Il Nuovo Cimento*, **102 B**, 471.

Press, W.H., Flannery, B.P., Teukolsky, S.A. & Vetterling, W.T., 1988. Numerical Recipes, Cambridge University Press, New York.

Robertson, David I., 1990, Investigations and developments relevant to a prototype laser interferometric gravitational wave detector. *PhD thesis*, Universtiy of Glasgow.

Robertson, David I., Private communication. 1991. Subject: Glasgow timing.

Robertson, David I., private communication. 1992. Subject: housekeeping data. At the Cardiff GW data analysis meeting, Easter 1992.

Rüdiger, A, 1990. Technical Report, Report on the Garching detector during the 100 Hour Run, Max Planck Institut für Quantenoptik, Garching, Germany.

Rüdiger, A., Private communication. 1992. Subject: general.

Rüdiger, A., Private communication. 1993. Subject: A/D sawtooth.

Sathyaprakash, B.S. & Dhurandhar, S.V., 1991, Gravitational wave data analysis of the coalescing binary filter: A criterion for choosing filters. *Physical Review D*, **44**, 3819.

Schilling, Roland, Private communication. 1991. Subject: Garching timing.

Schutz, Bernard F., 1985. A first course in general relativity, Cambridge University Press, Cambridge.

Schutz, B.F., 1986, Determining the Hubble Constant from gravitational wave observations. *Nature*, **323**, 310.

Schutz, B.F., 1988a. Data Analysis Requirements of Networks of detectors, In: *Gravitational wave data analysis: proceedings of the NATO advanced research workshop held at Dyffryn House, St. Nicholas, Cardiff, Wales, 6-9 July 1987*, p. 315, ed. Schutz, B.F., Kluwer Academic Publishers, Dordrecht.

Schutz, B.F., 1988b. Sources of gravitational radiation, In: *Gravitational wave data analysis: proceedings of the NATO advanced research workshop held at Dyffryn House, St.*

Nicholas, Cardiff, Wales, 6-9 July 1987, p. 73, ed. Schutz, B.F., Kluwer Academic Publishers, Dordrecht.

Schutz, B.F., 1989. Gravitational Radiation, In: *Proceedings of the Fourteenth Texas Symposium on Relativistic Astrophysics*, 27, Annals of the New York Academy of Sciences.

Schutz, Bernard F., 1991. Data Processing, Analysis, and Storage for Interferometric Detectors, In: *The Detection of Gravitational Waves*, p. 406, ed. Blair, D.G., Cambridge University Press.

Taylor, J.H. & Weisberg, J.M., 1989, Further experimental tests of relativistic gravity using the binary pulsar PSR 1913+16. *Astrophysical Journal*, **345**, 434.

Thorne, K.S., 1987. Gravitational Radiation, In: *300 Years of Gravitation*, p. 330, eds Hawking, Stephen & Israel, Werner, Cambridge University Press.

Thorne, K.S., 1992, On Joseph Weber's new cross-section for resonant-bar gravitational-wave detectors. *Recent Advances in General Relativity*, 4, 241. Essays in honour of Ted Newman.

Tifft, W.G., 1976, Discrete states of redshift and galaxy dynamics. *Astrophys. J.*, **206**, 38.

Trimble, V., 1988, 1987A — The greatest supernova since Kepler. *Reviews of Modern Physics*, **60**, 859.

Watkins, W.J., 1991, A Prototype System for Gravitational Wave Data Analysis. *PhD thesis*, University of Wales.

Weber, J., 1960, Detection and generation of gravitational waves. *Physical Review*, **117**, 306.

Weber, J., 1984. *Foundations of Physics*, **14**, 12.

Zucker, M., Group communication. 1992. Subject: glitch.