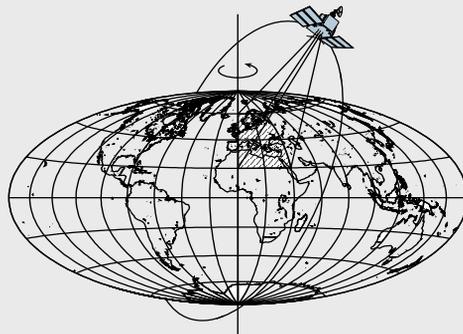


Small Anomalous Mass Detection from Airborne Gradiometry

by

Puttipol Dumrongchai



Report No. 482

Geodetic Science and Surveying

The Ohio State University
Columbus, Ohio 43210

March 2007

**SMALL ANOMALOUS MASS DETECTION FROM AIRBORNE
GRADIOMETRY**

By

Puttipol Dumrongchai

Report No. 482

**Geodetic Science and Surveying
School of Earth Sciences
The Ohio State University**

March 2007

ABSTRACT

A new generation of gradiometer technology is currently under development based on atom interferometry and applicable to ground and airborne mapping of geologic or anthropogenic features with signal strength as low as a few Eötvös, entirely embedded in noise and geological background. With high sensitivities of future airborne gradiometers, it may be possible to detect such anomalous sources with careful data processing. Both the detection and the estimation of parameters of the feature can be solved as an inverse problem in potential theory. However, one can also use methods developed in communications theory, provided one has some a priori, possibly uncertain knowledge of the feature in question. We constructed a matched filter as well as a sophisticated estimation technique to detect and characterize particular small mass anomalies within general geologic background noise using individual gradient and six gradient combination measurements at low aircraft/helicopter altitudes of ranges of 10-30m above terrain clearance. Since both detection and estimation portions requires the inversion of large sizes of covariance matrices, we applied an orthogonal transformation to the matrices, which become diagonal and can then be easily inverted. In addition, the performance of the detection and estimation procedures is quantified by standard test statistics. With these tests, probabilities of false alarm and detection may be assigned to the detection results. We present numerical results in different noise circumstances, for instance, a simulation of airborne gradiometry over moderate terrain with the inclusion of $1\text{E}/\sqrt{\text{Hz}}$ instrumental white noise. The proposed approaches are explored and evaluated for their effectiveness in association with location, orientation, size, and depth of a mass anomaly, and in the use of power spectral density (psd) models versus empirical psd's obtained from the noise backgrounds. The numerical results show that a small anomaly, e.g., 2m x 2m x 10m, is detectable at shallow depths by an appropriate matched filter using, not only the empirical psd's and the gradient component Γ_{33} , but also the psd models and the six-gradient combination. However, the analysis shows that a strong noise level, low spatial resolution, and unknown depth limit the anomaly detectability. The parameter estimation performed through an iterative least-squares process was shown to be successful in estimating locations, orientations, and depth of the anomaly. Hypothesis testing by means of the F-test was used to quantify the performance of the estimation process.

PREFACE

This report was prepared by Puttipol Dumrongchai, a graduate student, Division of Geodesy and Geospatial Science, School of Earth Sciences at the Ohio State University under the supervision of Prof. Christopher Jekeli.

This report was also submitted to the Graduate School of the Ohio State University as a dissertation in partial fulfillment of the requirements of the Ph.D. degree.

ACKNOWLEDGMENTS

I wish to express my sincere gratitude to my adviser Professor Christopher Jekeli, for intellectual support and guidance throughout the development of this research.

My appreciation goes to Professor Burkhard Schaffrin for giving constructive and valuable suggestions in the concept of parameter estimation during his sabbatical leave.

I want to acknowledge Professor C.K. Shum and Professor Michael Bevis for review this study and being a part of the evaluation committee.

I would like to thank Karla Edwards and Irene Tesfai for their friendship. I also want to thank Royal Thai Government for financial support through the Ph.D. program until the beginning of the year 2003 (before this research had been started).

My deepest gratitude must go to my mother, Mala Dumrongchai, and my special nephew, Kuntaratana Dumrongchai, for their support and patience. Their love and encouragement are always with me. Without them, especially, at a critical time in my life at the death of my brother, Angoon Dumrongchai, in September 2004, this work would not have been possible.

TABLE OF CONTENTS

	Pages
Abstract.....	ii
Dedication.....	iii
Preface.....	iv
Acknowledgments.....	v
List of Tables.....	ix
List of Figures.....	xiii
Chapters	
1. Introduction to small anomaly mass detection.....	1
1.1 Introduction.....	1
1.2 Reviews of previous and current works.....	3
1.3 A preliminary effort to apply detection and estimation theory for small anomaly detection.....	5
1.4 The proposed research.....	8
1.5 Organization.....	10
2. The potential theory of gravitational gradients and detectability.....	11
2.1 Introduction.....	11
2.2 The potential theory of gravitational gradients.....	12
2.3 Gravitational gradients of geometric forms.....	14
2.4 Transformation.....	17
3. The theory of signal detection by using matched filters.....	19
3.1 Introduction.....	19
3.2 Signal detection.....	19
3.3 Definition of a matched filter.....	20
3.4 The matched filter for correlated noise.....	22
3.5 Hypothesis testing under Neyman-Pearson criterion.....	24

3.6 Performance of detection.....	26
3.6.1 White noise case.....	27
3.6.2 Correlated noise case.....	29
4. The optimal combination of six components of gravitational gradients.....	33
4.1 Introduction.....	33
4.2 The combination of matched filters.....	33
4.3 Hypothesis testing and performance of detection.....	38
5. Numerical discussions on shallow small anomalous mass detection.....	40
5.1 The simulated survey of airborne gradiometry.....	40
5.2 CASE I: White noise environment.....	47
5.3 CASE II: Correlated noise environment.....	54
5.4 CASE III: Anomaly detection in correlated noise using the combination of six gradient components.....	71
5.5 CASE IV: Anomaly detectability versus sizes and depths.....	81
5.6 CASE V: Anomaly detectability versus orientations.....	94
5.7 Performance of matched filters.....	102
6. The theory of parameter estimation.....	109
6.1 Introduction.....	109
6.2 The random effects model and hypothesis testing.....	110
6.3 Implementation	117
6.3.1 Single data track and single sensor (SS).....	117
6.3.2 Single data track and multiple sensors (SM).....	120
6.3.3 Multiple data tracks and single sensor (MS).....	124
6.3.4 Multiple data tracks and multiple sensors (MM).....	127
7. Numerical discussions on parameter estimation.....	131
7.1 Introduction.....	131
7.2 CASE I: The impact of the a priori covariance matrix on localization and test statistic.....	134
7.3 CASE II: Predicted location and fixed orientation using individual gradient components in $0.01E^2/\text{Hz}$ (or $3E^2/(\text{cyc}/\text{m})^2$) white noise environment.....	136
7.4 CASE III: Predicted location and fixed orientation using individual gradient components in $1E^2/\text{Hz}$ (or $300E^2/(\text{cyc}/\text{m})^2$) white noise environment.....	146
7.5 CASE IV: Predicted location and fixed orientation using the combination of all six gradient components in $0.01E^2/\text{Hz}$ (or $3E^2/(\text{cyc}/\text{m})^2$) white noise environment.....	155
7.6 CASE V: Predicted location and fixed orientation using the combination of all six gradient components in $1E^2/\text{Hz}$ (or $300E^2/(\text{cyc}/\text{m})^2$) white noise environment.....	158

7.7 CASE VI: Predicted location and fixed orientation using individual gradients in geological background.....	162
7.8 CASE VII: Predicted location and fixed orientation using six combination of gradient components in geological background.....	169
7.9 CASE VIII: Predicted location and orientation in $0.01E^2/\text{Hz}$ white noise environment using six gradient components.....	172
7.10CASE IX: Estimation of density contrast.....	176
8. Conclusions and future work.....	178
References.....	181
Appendix A. The reciprocal distance power spectral density models of gravitational gradients.....	188
Appendix B. A rectangular prism and its partial derivatives with respect to origin and orientation.....	190
B.1 A rectangular prism.....	190
B.2 Origin and orientation.....	191
B.3 The partial derivatives of a rectangular prism in the frequency domain.....	196
Appendix C. Rearrangement of data tracks by applying orthogonal transformation.....	197
C.1 Single data track and single sensor (SS).....	197
C.2 Single data track and multiple sensors (SM).....	199
C.3 Multiple data tracks and single sensor (MS).....	203
C.4 Multiple data tracks and multiple sensors (MM).....	205

CHAPTER 1

INTRODUCTION TO SMALL ANOMALY MASS DETECTION

1.1 Introduction

In the early 1980's, the US Air Force, NGA (then Defense Mapping Agency), and Bell Aerospace Corp. developed a new gravity gradiometer system, called the Gravity Gradiometer Survey System (GGSS) that, in conjunction with the Global Positioning System (GPS), was designed to survey Earth's gravity field on a regional basis (Jekeli, 1987 and 1993). Although the GGSS was successfully tested in a limited survey of a large area, it was unfortunately terminated due to lack of funds. The existence of the GGSS has drawn geophysicists' attention towards the use of airborne gravity gradient data not only for mapping the earth's gravity field, but also for mineral exploration and other small anomalous mass detection.

According to Jekeli (1993), the practical accuracy of the GGSS data is on the order of 10E (1E = 1Eötvös= 10^{-9} sec⁻²) per ten-second average (corresponding to noise performance of $30E/\sqrt{\text{Hz}}$). The data were obtained on a grid of orthogonal tracks spaced 5 km apart, at an altitude of about 700 m in the Texas/Oklahoma area. In 1999, BHP's Project Falcon and Lockheed-Martin introduced the world's first airborne gravity gradiometer (AGG) system for mineral exploration with noise performance as small as $7E/\sqrt{\text{Hz}}$ (van Leeuwen, 1999; Lee, 2001). The AGG system was capable of sensing small variations in the Earth's gravity at the level required in mineral exploration with 100m resolution at a typical altitude of 100m above ground. Lui et al. (2001) demonstrated the detectability of Kimberlite pipes as small as 100m in diameter.

Despite the AGG's satisfactory outcome, detecting finer and smaller underground structures whose gradient signals have much shorter wavelengths, normally located close to the ground surface, is difficult. Romaides et al. (2001), for instance, conducted a ground gradiometric survey in the static mode over an underground Missile Alert Facility (MAF-EO), located at Vandenberg Air Force Base, California, by using a vertical spin-axis Lockheed Martin gradiometer. The gravity gradient measurements indicated a peak gradient signal of approximately 30E over the buried facility, whose center lies at 12m depth below the ground surface. Kastella (2003) also studied the possibilities of using a precise airborne gradiometer to detect small objects. Nonetheless, his assessment of flight levels is still unclear. Although, a precise gradiometer system is essential for detecting small objects, to date, no current technology has succeeded in yielding a specification of sensing 1E gravity gradient or less for resolutions down to 20-40m.

Meanwhile, the Stanford research group has developed the cold atom interferometry gradiometer (Snadden et al., 1998; Peters et al., 2001; McGuirk et al., 2002). The overall sensitivity of noise performance is $4E/\sqrt{\text{Hz}}$ in laboratory tests. Also, the research team at the University of Western Australia (UWA) has developed the OQR¹ gravity gradiometer, which has achieved its target noise performance of close to $1E/\sqrt{\text{Hz}}$ in laboratory tests (van Kann, 1992; Matthews, 2002). With hope towards the success of those instrumental developments in the near future, it is expected that the new generations of precise gradiometers will be capable of detecting 2m-width x 2m-height x 1000m-length through 2m-width x 2m-height x 10m-length mass anomalies at depths of 10-50m, potentially at a very close range of 30-100m.

In general, gravity gradiometry is the measurement of the spatial gradients of the gravity vector. It is a method used to extract physical information of the Earth's upper crust for determination of the gravity field and also to explore mineral resources and oil deposits. Unlike gravimetry on a moving platform, airborne gradiometry is not dependent on linear accelerations of the vehicle and only sensitive to angular accelerations, which can be measured by gyroscopes. The system may consist of a number of gradient sensors permitting the determination of the five independent components of the gravity gradient tensor. The gradiometer is capable of sensing short-wavelengths of the Earth's gravity field, generated by the mass density of the upper crust and topography. It may be able to detect and estimate small anomalous masses near the Earth's surface. Another attraction is that each gradient component itself intrinsically provides a different aspect of the geometric shape of the anomaly. This is because the gradient tensors are related to spatial directions and thereby are influenced by shapes of the mass sources. However, it is difficult to detect high-resolution density anomalies due to the complexity of the instrument, the angular acceleration environment of the vehicle, and the sensitivity of the instrument to the very near field. Moreover, the varying terrain itself may obscure the comparatively small signals produced by sub-surface mass anomalies. Even today, with high instrumental sensitivities, processing the data with additive noise is a major challenge among geodesists and geophysicists because it fundamentally is a geophysical inverse problem which has inherent non-unique solutions. Robust signal processing methods need to be developed to extract meaningful signals from the geographical background variation.

According to Jordan (1978), over 99% of the gravity gradient signals are due to the visible topography as well as the underlying mass anomalies within 1 km below the geoid. This implies that the signals correspond to a contribution of terrain mass. One basic idea of detecting mass anomalies is that, by subtracting known or theoretical gradients from the observed gradients, the differences essentially yield the signature of anomalies created by unknown density contrasts. The theoretical gradients can be computed using forward modeling approaches (e.g. flat-top rectangular prism and/or slope-top triangular prism) applied to Digital Elevation Models (DEM) with the assumption of a constant density. Such finite element models significantly depend on the

¹ The Orthogonal Quadrupole Responder (OQR) is designed to measure the off-diagonal components of the gravity tensor using two quadrupole responders that are mounted orthogonally to each other.

type of geometrical representation and grid spacing, which can corrupt the information to be extracted.

This research will focus on one class of methods of signal detection and parameter estimation to deal with the inverse problem to determine horizontal location, density contrasts, and depths of small mass anomalies. Because the gravity gradients have a local behavior, the signals themselves can be severely corrupted by the effect of nearby topography. Filtering techniques for the detection problem will be developed on the basis of the matched filter. The random effects model including stochastic linear hypotheses (Schaffrin, 1987, 1989, and 2001; Schaffrin and Bock, 1994) will be adapted to the problem of parameter estimation. Due to the fact that real data are not available, the appropriate surveys of airborne gradiometry will be simulated based on the combination of the 1987 GGSS data set and finite elements of flat-top rectangular prisms from 1"x1" DEM (Smith et al., 2001), provided by the National Geodetic Survey (NGS).

1.2 Review of previous and current works

For over half of a century, several sophisticated techniques for geological interpretation have been developed for magnetic and gravity data, but only a few have been applied to gravity gradient data. After Lancaster-Jones (1932), Nettleton (1976), Lorenzini et al. (1988), Paik and Moody (1993), van Kann et al. (1993), and the development of the GGSS (Jekeli, 1987 and 1993), the usage of gravity gradient has become increasingly interesting due to its possibilities for such applications ranging from geophysical explorations to space gravity mapping. Klingele et al. (1991) showed that the use of vertical gravity gradient is capable of extracting detailed information of the Earth's structures superior to that of gravity data. The investigation was done by means of the Euler deconvolution method (Thomson, 1982). Marson and Klingele (1993) further investigated the advantage of using the gradient signal to delineate density discontinuities in three dimensions.

Thurston and Smith (1997) and Smith et al. (1998) introduced the improved Source Parameter ImagingTM (iSPITM) method for depth determination, which is independent of geometric shapes assigned to the causative bodies. Aside from the methods of depth estimation, vertical gradient data were utilized to model the Earth's crustal density distribution (Forsberg, 1984). However, because, in theory, an infinity of mass distributions solve such an inverse problem, such modeling is not robust without constraints. Vasco (1989) proposed a sophisticated way to deal with the ill-posed condition by introducing inequality constraints in linear and nonlinear inversion problems over the space of finite rectangular elements. Then, he compared the capability of airborne gradiometer data with gravity data, which were gathered from the airborne survey, conducted by the GGSS over Texas/Oklahoma area. The result showed that the gradiometer data improved the model parameter resolution and in turn yielded a maximum density error of 0.1 g/cm^3 above 20km depth.

Vasco and Taylor (1991) extensively tested the nonlinear problem using the same set of gradient data to determine the basement topography with a spatial resolution of about 1.9 km. The solution resulted in mapping the basement topography as deep as 12km. Talwani (2003) further developed a joint inverse method based on the nonlinear problem under a constraining condition (e.g., either fixed tops or bases of rectangular

prisms). The method used gravity data and the full gradient tensor on a 100m grid in a small area of size 3km x 3km. The author showed a significant improvement in obtaining positions of the top and the base of salt domes at 1km and 3km depths. The misclosures of modeled and observed gradient and gravity values varied in the range of 2-5E and 0.5-1mgal, respectively. A crucial suggestion was made by Tawani (ibid.) that with 5E sensitivity of a moving-base gradiometer, a target at a depth of about 600m was detectable if its wavelength is as small as one kilometer and the flight height is about 100m.

Another work by Zhdanov et al. (2004) supports Talwani's suggestion. They developed a new method for interpreting the full gradient tensor based on three-dimensional regularized focusing inversion. The gradient data used in this research came from FalconTM AGG survey over the Cannington Ag-Pb-Zn orebody (size of about 400m² at depths varying from 50m to 500m below the surface) in Queensland, Australia. The processed data corresponded to an effective sensor height of 120 m above ground level with sampling approximately 20 m along survey lines and with 100 m tracking spacing. The authors showed the potential of the proposed method to reconstruct a sharper image of the geological target as deep as 500m.

One of major concerns in airborne gradiometry is the effect of topographic mass, becoming more significant at decreasing altitude. A correction to measure gradients must be applied in order to reduce the effect as much as possible. Nonetheless, there are two questions that arise in the numerical process of the terrain correction. First, how can the complexity of the terrain be modeled with sufficient accuracy? Second, how much knowledge of the terrain mass densities is required? These concerns have not yet been fully investigated and are outside the scope of this study.

There are studies by different authors (e.g., Chinnery, 1961; Dorman and Lewis, 1974; Hammer, 1976), which are related to the methods of terrain correction with a constant density. However, these studies did not consider the terrain effects in areas of complex topography. Further analyses were performed in a very rough terrain along the Rocky Mountain Chain by Tziavos et al. (1988), who developed an algorithm using the Fast Fourier Transform (FFT), the definition of which can be found in Brigham (1988). Then, Tziavos et al. (1988) compared FFT gradient computation with that generated by finite rectangular prisms. It was shown that the terrain effect significantly depended on the ratio between flight elevation and grid spacing, the type of terrain representation, the attenuation of the gradient field at flight level, and the order of expansion of the terrain correction integral based on Parker's formula (Parker, 1972 and 1994). The authors concluded that obtaining accuracy of 0.5E for a flight height of 0.6 km above ground required a grid spacing of about 0.25km. However, most of these analyses did not consider smaller scales of grid spacing (e.g. less than 100-meter levels).

Using a one-arcsecond DEM, Jekeli et al. (2003) computed the full tensor of gravitational gradients from slope-topped triangular and flat-topped rectangular prisms at low altitudes of 10m and 100m. The numerical differences in magnitude increase up to 9E and 37E at 100-m height and 10-m height, respectively. These studies agreed with previous analyses, given that the gradients strongly depend on the height of computation. Further work was done by Jekeli and Zhu (2006), who showed that the resulting gradients from triangular prism models depends on the diagonalization of the rectangular

grid. The authors also compared different modeling methods such as triangular prism formulas, rectangular prism/numerical integration formulas, Forsberg FFT, and Parker FFT. However, more studies are still needed to better understand the consequences of terrain representation effects towards data processing.

One may also consider the downward continuation of gradient data using, for example, least-square collocation (Moritz, 1980) or Poisson's integral (Heiskanen and Moritz, 1979). A sophisticated technique of downward continuation ideally should minimize noise levels inherited from various data sources in some optimal fashion. White and Goldstein (1984) developed a continuation technique combining minimum-variance estimation and computational efficiency by performing some data processing in a suboptimal manner. Jekeli (1985) further studied White and Goldstein's method by comparing with an optimal procedure such as least squares collocation. He showed the additional advantage of yielding error estimates by a suboptimal technique exemplified by frequency domain collocation. Other downward continuation techniques can be found in Jekeli (1986), Gleason (1988), and Wang (1988), who however applied them to estimate gravity quantities from gradient data. A number of numerical investigations are still needed to broaden the scope of techniques applicable to precise airborne gradiometry.

Another aspect to consider in the detection and estimation problem is measurement noise. Filtering techniques should minimize the noise and in turn combine the available gradient components in an optimal fashion because the gradient signals are comparatively small and could be masked by the effect of topography and the longer wavelength signals. One of the robust filtering techniques, widely used in geophysical signal processing, is the Wiener filter (Wiener, 1949; Gunn, 1975 and 1978), which is based on the principle of minimum-mean-square error. Pawlowski and Hansen (1990) introduced the technique to gravity anomaly separation and showed that the superiority of Wiener filter's scheme over conventional band-pass filters significantly depended on utilizing accurate geologic information. However, the filter did not perform well in the presence of increasing spectral overlap between the gravity signal and the spectral contributions of other geologic sources. One possible explanation is the inability of the filter to reconstruct meaningful frequencies where the signal-to-noise ratio is low. The matched filter is a robust technique that should be considered because it corresponds to a filter whose chosen filter coefficients yield maximum signal-to-noise ratio in the output (Turin, 1960 and 1976; Poor, 1983; Chen and Kassam, 1985; Cadzow, 1987). This filter is relatively new in the field of gravimetric geodesy and will be the focus in this research.

1.3 A preliminary effort to apply detection and estimation theory for small anomaly detection

Although a new generation of precise gradiometers is destined to in the civilian community in the near future, the detection of small anomalous masses still has challenges with respect to signal processing capability. Some crucial factors are emphasized as follows:

- The six components of the gravitational gradient tensor generated by the anomalies are small compared to the geological background,

- All wavelength filtering is limited in efficiency if there exist overlapping spectra between the signals and geological background,
- There is rapid signal attenuation with increasing flight altitudes,
- Aliasing effects are due to data spacing related to flight speed,
- It is difficult to identify the anomalies if measurements are corrupted by noise,
- Detection by visualization may be difficult under some circumstances,
- The spatial location of the detected anomaly may not be as close as possible to its true location. An optimal estimator is needed to refine the spatial location.

One of the preliminary investigations by Dumrongchai and Jekeli (2004a,b) applied the concept of the matched-filter technique to better understand the above factors that influence the gravitational gradients. The matched filter (MF) was first introduced by North (1943) in the field of communication systems. Since then, it has been widely used in various applications of electrical engineering, especially, in radar detection and pattern-recognition. The development of the MF has other applications in signal detection. For example, Poor (1983) and Geraniotis and Poor (1987) proposed robust matched filters for optical receivers. Although the MF has a long history of development, the filter has rarely appeared in the field of physical geodesy. One comprehensive work was done by White et al. (1983), who introduced MF for the detection of characteristic geoid undulation signatures of seamounts in SEASAT radar altimeter data. Nonetheless, applications of the matched filter related to mass anomaly detection remains untested in geodesy and geophysics.

Dumrongchai and Jekeli (2004a) developed the matched filter for the gravitational gradients and preliminarily tested the filters with simulated components of the gradient tensor using DEM (Jekeli and Zhu (2004), personal communication). Figure 1.1 shows the results of matched filtering for the case of a non-random effect s_{ij} and white Gaussian noise. Based on the filtering results, the authors made the following conclusions. Firstly, the ability of matched filters to locate the anomaly is significantly impeded not only by noise but also by spatial sampling. Secondly, since matched filters require knowledge of the signal gradients, they perform poorly when geological noise spectra overlap the spectrum of the anomaly. Finally, the geological background should be removed beforehand to enhance the capabilities of matched filtering.

The problem of detecting the gravity gradients due to shallow small mass anomalies in noise can be envisaged as extracting a very small signature from that total gravity gradient field of the Earth; see Figure 1.2. The gravitational gradients due to small mass anomalies are relatively small and span small ranges of very short wavelengths lying between 10m and a few kilometers, compared with much longer wavelengths of the total gradient field of the Earth.

The total gradient field measured by a gradiometer can be differenced with a gradient field model containing long, medium, and some short wavelengths, so that any residual in the measurement will correspond to unmodelled anomalies, potentially the anomalies in question. Nonetheless, the wanted signals due to the anomalies can be distorted by measurement noise. The outcome of such combinations can be described as the observation space.

We look for a particular signal due to one or more target anomaly and attempt to decide if the target is in observation space. Hypotheses testing based on a decision rule is applied to the problem where H_0 corresponds to no target and H_i (with $i = 1, 2, 3, \dots$) corresponds to the presence of the i^{th} target; i equals one if there is only one target of interest. The essential idea is to model the gradient signal due to the target anomaly if we have complete knowledge of its signal waveform, i.e., we are able to construct its corresponding matched filter. To decide which hypothesis is true, the designed matched filter is applied to aid the decision. For several anomalies in question, a bank of matched filters should be used. In practice, it is common that several interfering sources, e.g., masses of local topography and near surface density anomalies contribute to the filter output.

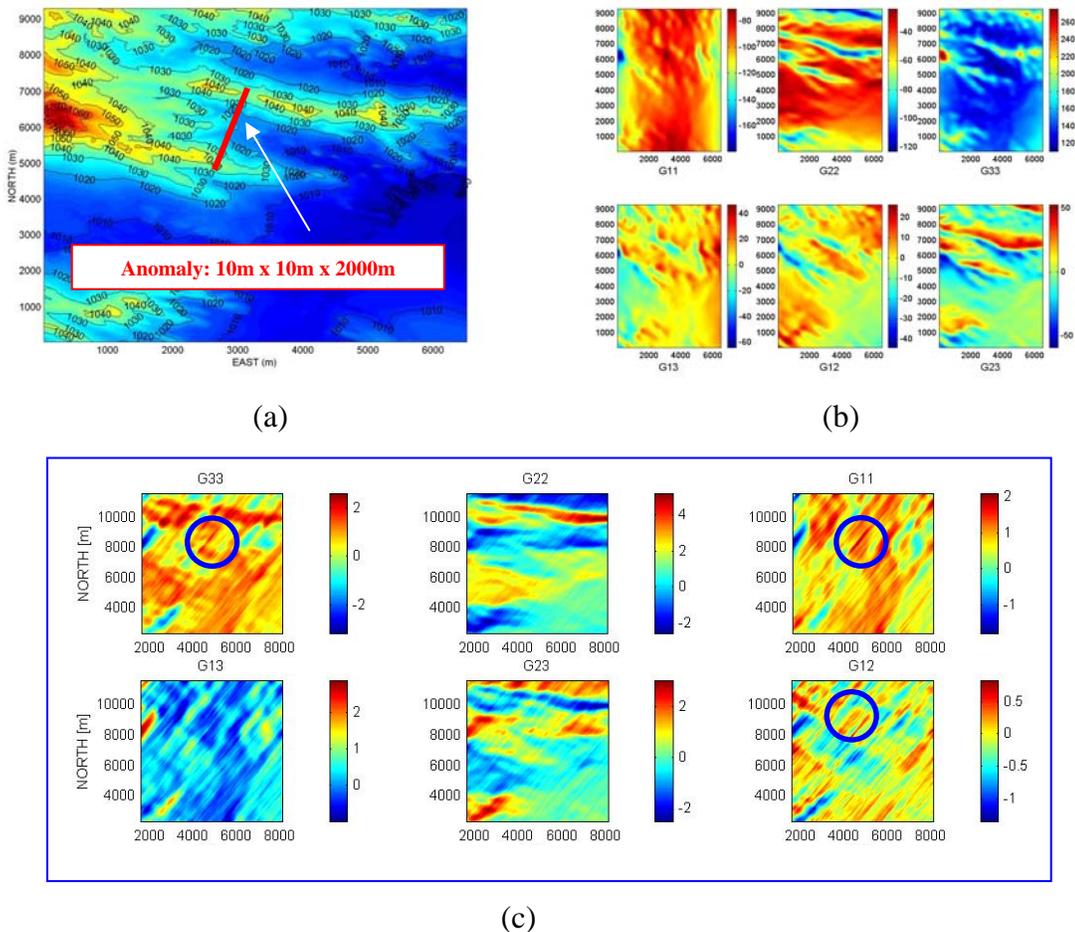


Figure 1.1: (a) One-arc second DEM and detection of a small anomaly (10m x 10m x 2000m) in Montana/Wyoming area, (b) six gravitational gradient components of the topography and the small anomaly generated using Fast Fourier Transform approach, and (c) matched filter outputs of the gradients--circle indicates the successful location of the small anomaly (no successful location for Γ_{22} , Γ_{13} , and Γ_{23}).

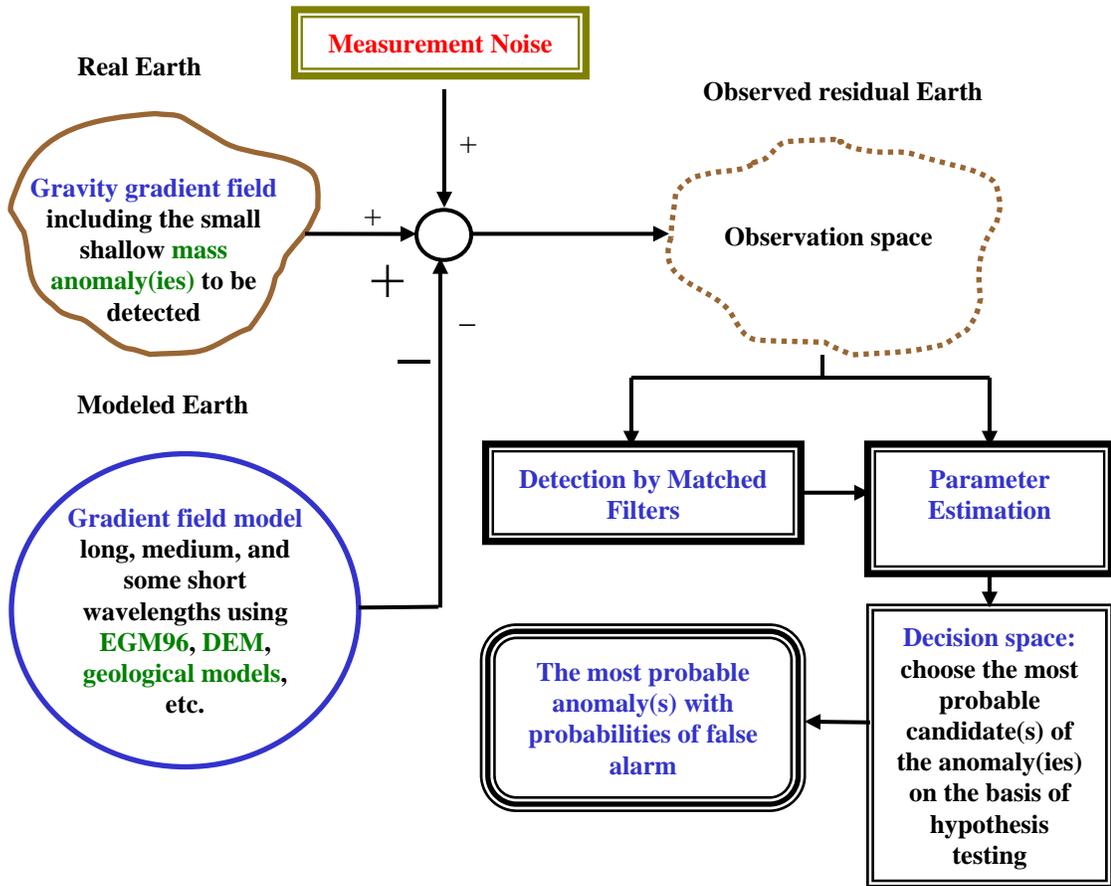


Figure 1.2: Signal detection and parameter estimation for small anomaly detection.

Moreover, we may not have much knowledge about the signal's waveform. These factors significantly degrade the performance of matched filtering. Therefore, the idea of parameter estimation becomes important to aid in the decision of choosing the most probable candidate of the target anomaly.

1.4 The proposed research

Gravity gradiometry with its achievable measurement sensitivities provides an enormous contribution to the problem of detecting small shallow anomalous mass. Overall, the study herein will directly aims at (1) the simulation of gradiometric airborne surveys based on the combination of the 1987 GGSS data set and finite elements of flat-top rectangular prisms, (2) the development of multiple sets of the gradient matched filters for the optimal hybrid matched filter of six gradient components, (3) signal detection and parameter estimation methodology, and (4) comparative assessments of capabilities of the combination techniques and data requirements to achieve the ultimate goal of small mass anomaly detection and estimation.

Since a precise gradiometer that measures all components of the gravitational gradient tensor is under development, with overall noise performance of about $1 \text{ E}/\sqrt{\text{Hz}}$, actual gradiometric data with such high precision are not yet available. For this work, six components of the gravitational gradient tensor will be simulated by using the 1987 GGSS data set and DEM.

The matched filter is developed for the six gravitational gradient components individually, as well as for combinations of gradients. Each component of the gravitational gradient tensor describes the local structure of the gradient field and represents particular characteristics at the same location where measurements are obtained by the gradiometer. To enhance the capabilities of matched filtering, the idea of a vector random process is considered. The hybrid-matched filter optimally combines six gradient components in terms of the noise covariances between different types of components. Furthermore, multiple sets of matched filters are also developed. The idea of it is to select the filter for each gradient component, which gives the highest signal-to-noise-ratio output exceeding a given threshold.

Throughout the processes stated above, in order for the matched filter to be optimal, complete knowledge (except location) is required about the mass anomaly to be detected. In practice, we vaguely know about the anomaly. However, theoretically at least, the structures of the gravitational gradients follow the Newtonian law of gravitational potential and can be described as deterministic functions associated with unknown parameters such as density contrast, location, orientation, and so on. These attributes are closely related to the problem of parameter estimation which is expanded to include signal detection. A detection and parameter estimation approach is proposed and all six components of the gradient tensor will be considered, as shown in Figure 1.3. The performance of detection and estimation will be assessed through the test statistic in view of the probability of miss (POM) and the probability of false alarm (POF). The test statistic will measure the quality of the proposed system. In other words, it will indicate the capabilities of that system to discriminate the gradient signals due to an anomalous mass from the geological noise background.

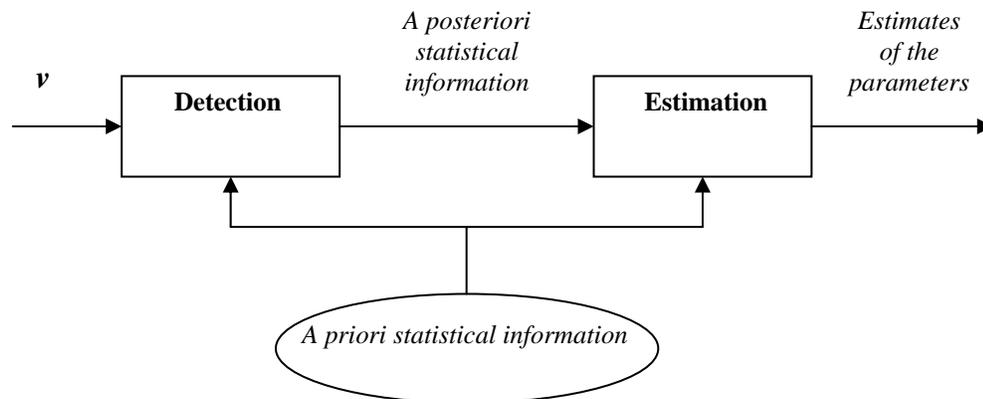


Figure 1.3: Detection and parameter estimation.

In conclusion, the ultimate goal of this study is to develop and analyze signal detection and parameter estimation techniques using all components of the gravitational gradient data (from a simulated survey of airborne gradiometry) to determine the horizontal location, depth, orientation, and also density contrast of small mass anomalies. Each gradient component as well as combinations of components will be assessed to determine their usefulness.

1.5 Organization

This dissertation is organized in the following manner. In Chapter 2, we study the potential theory of gravitational gradients and detectability is studied. Then, in Chapter 3, we move to the fundamental theory of signal detection from the perspective of communications theory and apply its concept to the problem of small mass anomaly detection.

The optimal combinations of six components of the gravitational gradient tensor are developed in Chapter 4. In Chapter 5, a prospective airborne gradiometric survey is simulated. Then, using the simulation, developments in of Chapters 3 through 4 are tested and assessed in order to determine their mass anomaly signal-detection capabilities, in association with various combinations of observed gradient components.

Chapter 6 is dedicated to parameter estimation. The random effects model and hypothesis testing are discussed and are developed to support the problem of small mass detection. In Chapter 7, numerical examples are discussed to show the robustness of the method. Finally, in Chapter 8, we conclude our results and suggest further work that can be focused to enhance the capability of detection by matched filters and parameter estimation.

CHAPTER 2

THE POTENTIAL THEORY OF GRAVITATIONAL GRADIENTS AND DETECTABILITY

2.1 Introduction

According to Nettleton (1976), gravity gradiometry was one of the earliest methods used to explore for oil deposits. The first gradiometer used for this purpose was the torsion balance, invented by Lorand von Eötvös in Hungary around the turn of the last century. Although exquisitely sensitive (for its time) to local crustal density variations, it was cumbersome to operate and was soon replaced by the scalar gravimeter that measures the (relative) magnitude of gravity acceleration. Static gravity as a whole, whether with gravimeters or gradiometers, is labor intensive and difficult to carry out in remote areas of limited accessibility. Thus, airborne gravimetry is viewed as the most efficient method to map the near-Earth gravity field. However, the main difficulty with a moving platform is the inseparability of specific forces coupled with gravitational accelerations to produce the total (kinematic) accelerations of the vehicle. On the other hand, airborne gradiometry using a gravity gradiometer on a platform with a gyroscope is believed to yield gravity maps of the near-Earth geopotential field without the need to determine the kinematic linear acceleration of the vehicle.

In general, modern gradiometers are sufficiently sensitive to the point that the integration time to produce an accurate measurement is relatively short, meaning that the resolution on a moving vehicle is high (on the order of hundreds of meters) and maybe higher (on the order of tens of meters). The resolution (by gradiometer) usually is much higher than that of a gravimeter (normally on the order of several, even tens of kilometers). Combined with the high spatial resolution, the high accuracy of a gradiometer implies that it is able to sense the fine structure of the gradient field due to small shallow mass anomalies. In the intelligence community, gradiometers are considered as a means to identify the small anomalies that inherently have sharp density contrasts relative to the surrounding environment. The detection of these anomalies is enhanced considerably by the ability to measure the gradient (or directional derivative) of the corresponding gravitational signal. And, since gradiometers can be combined to sense all six gradients, the ability to characterize such anomalies is further maximized.

Finding shallow mass anomalies stands out as one of the toughest technical challenges in the intelligence community's efforts to locate and assess targets of interest anywhere inside the Earth (Streland, 2003). To find and target the anomalies effectively, it is necessary to understand how they are built. Understanding the nature of the anomalies provides valuable information of what to look for in the problem of detection.

This chapter, focuses on the fundamental theory of gravitational gradients. Detectability of the anomalies with respect to spatial resolution (half-wavelength) and flight altitudes at which the measurement are obtained are discussed. The technical details concerning instrumental developments will not be discussed any further in this work, and can be found in Jekeli (1987), Peters et al. (2001), Matthew (2002), and McGuirk et al. (2002).

2.2 The potential theory of gravitational gradients

Starting from Newtonian potential theory¹, the gravitational (not gravity) potential² (a scalar, zero-order-tensor function), U , at a point in a Cartesian coordinate system, $\mathbf{x} = (x_1, x_2, x_3)^T$, due to an (attracting) anomalous mass distribution having density function, $\rho(x'_1, x'_2, x'_3)$, and volume, v , is defined by:

$$U(\mathbf{x}) = k \iiint_v \frac{\rho(x'_1, x'_2, x'_3)}{r} dx'_3 dx'_2 dx'_1 \quad (2.1)$$

with the slant distance

$$r = \sqrt{(x_1 - x'_1)^2 + (x_2 - x'_2)^2 + (x_3 - x'_3)^2} \quad (2.2)$$

where k is the Newtonian gravitational constant and $\mathbf{x}' = (x'_1, x'_2, x'_3)^T$ is the integration point. The vector of gravitational acceleration, \mathbf{g} , is defined as the gradient of the potential (the first-order tensor):

$$\mathbf{g}(\mathbf{x}) = \nabla_x U(\mathbf{x}) = k \iiint_v \nabla_x \frac{\rho(x'_1, x'_2, x'_3)}{\ell} dx'_3 dx'_2 dx'_1 \quad (2.3)$$

where ∇_x denotes a vector of partial derivative operators, e.g.,

$$\nabla_x = \begin{bmatrix} \partial/\partial x_1 \\ \partial/\partial x_2 \\ \partial/\partial x_3 \end{bmatrix} \quad (2.4)$$

where the subscript identifies the variables with respect to which the differentiation is performed. The gradient of the gravitational acceleration is a tensor of second order, consisting of the partial derivatives of the gravitational vector components and arranged in matrix form:

¹ Understanding the concept of Newtonian potential theory is essentially needed for those who are unfamiliar with the field of physical geodesy. Readers are encouraged to read through some excellent textbooks such as Heiskanen and Moritz (1979), Torge (1989 and 2001), and Hofmann-Wellenhof and Moritz (2005).

² The *gravity* potential equals the *gravitational* potential *plus* the potential of *centrifugal* force.

$$s(\mathbf{x}) = \nabla_{\mathbf{x}} \mathbf{g}^T(\mathbf{x}) = \nabla_{\mathbf{x}} \nabla_{\mathbf{x}}^T U(\mathbf{x}) = k \iiint_V \nabla_{\mathbf{x}} \nabla_{\mathbf{x}}^T \frac{\rho(x'_1, x'_2, x'_3)}{r} dx'_3 dx'_2 dx'_1 \quad (2.5)$$

or

$$\mathbf{s} = [s_{ij}] \quad (2.6)$$

where the elements of the matrix are

$$s_{ij}(\mathbf{x}) = \frac{\partial^2 U(\mathbf{x})}{\partial x_i \partial x_j} \quad \text{with } i \text{ and } j = 1, 2, \text{ and } 3. \quad (2.7)$$

Clearly, the matrix of second partial derivatives is symmetric and Poisson's equation yields the constraint:

$$\nabla_{\mathbf{x}}^2 U(\mathbf{x}) = \nabla_{\mathbf{x}} \bullet \nabla_{\mathbf{x}} U(\mathbf{x}) = \text{trace}(\nabla_{\mathbf{x}} \nabla_{\mathbf{x}}^T U(\mathbf{x})) = -4\pi k \rho(\mathbf{x}) \quad (2.8)$$

If the points (i.e. \mathbf{x} 's) are outside the attracting mass, the constraint satisfies the Laplace's condition

$$\nabla_{\mathbf{x}}^2 U(\mathbf{x}) = s_{11} + s_{22} + s_{33} = 0 \quad (2.9)$$

Only five gradient components are independent, because of (2.8) and the symmetry property (i.e., $s_{ij} = s_{ji}$). As a consequence of how gradients were measured, for example, with the torsion balance, exploration geophysicists have defined combinations of the gradients as indicated below (Nettleton, 1976; Torge, (1989); Dransfield (1994)). Adopting a coordinate system in which the coordinates, x_1 and x_2 , are in the horizontal plane, the horizontal gravitational gradient vector determines the horizontal variation of the vertical component of gravitation, with amplitude and direction in the horizontal plane given by

$$s_H = \sqrt{s_{13}^2 + s_{23}^2} \quad \text{and} \quad \alpha_H = \tan^{-1} \left(\frac{s_{13}}{s_{23}} \right) \quad (2.10)$$

Another quantity is defined to describe the variation of curvature of the equipotential surfaces, $U = \text{constant}$, of the gravitational field. The curvature of a surface is the rate of change of the tangent to the surface and thus, for equipotential surfaces, it is related to the second derivatives of the potential. In particular, it can be shown that the difference between maximum and minimum curvature at a point is given by the scalar, s_C , called the differential curvature, where the minimum corresponds to the tangent in the (horizontal) direction, α_C , given by (Dransfield, 1994):

$$s_C = \sqrt{4s_{12}^2 + s_\Delta^2} \text{ with } \alpha_C = \frac{1}{2} \tan^{-1} \left(\frac{2s_{12}}{s_\Delta} \right) \quad (2.11)$$

and

$$s_\Delta = s_{11} - s_{12} \quad (2.12)$$

where s_Δ is called the inline gradient.

So far, the coordinate system being used for the entire study of small mass anomaly detection has not been specified. It is evident that the second derivatives of the kernel $1/r$ of (2.1) rapidly diminish the value of the integrand, when the integration points are a great distances from \mathbf{x} . As a consequence, airborne gradiometry is suitable for survey areas at local scales, whereas gravimetry is for regional scales. One may extend the integration over an equivalent plane, corresponding to planar approximation for the area of integration (i.e., the area of interest is considered using a flat Earth model). The local coordinate system (or frame³) can be legitimately defined to be associated with the North (N), East (E), and Down (D) directions. The subscripts (1,2,3) of point \mathbf{x} indicate (N,E,D) respectively.

2.3 Gravitational gradients of geometric forms

In general, the gravitational potential of (2.1) refers to the outcome of integration over the volume of an arbitrary shape of the attracting mass, whose density may vary at any point (i.e., it is an inhomogeneous body). It is impossible to determine a closed formula of an arbitrary body-type unless some assumptions are made. For simplicity, we assume that a simple geometric and homogenous body (or a finite sum of such bodies) can represent any anomalous mass. Simple geometric shapes delineating constant-density distributions, in fact, have dominated the modeling of crustal density anomalies, including the rectangular prism, the point mass (monopole) and the vertical cylinder and line segment, among others. The motivation for using these is always the mathematical simplicity with which the corresponding gradients can be formulated.

In this chapter, we are interested in the gravitational gradient (effect) due to the point mass and the rectangular prism with a constant density, whose closed formulas can be readily derived and also can be found in Forsberg (1984) and Jekeli et al. (2003). Other geometric bodies can be found in Dransfield (1994). According to Newton's law of gravitation⁴, the particular construction of the rectangular prism can be used to approximate, by compact aggregation, any given mass density distribution with a finite number of constant-density blocks (prisms) (Jekeli, 2002). The resolution and accuracy of the corresponding gradient are limited only by the number of such blocks. Some other geometric bodies can also be included. Numerical examples can be seen in Chapter 5. The analytic formulas of the gradient effects due to a point mass and a rectangular prism with a constant density, ρ , e.g., $\rho = 1.170 \text{ g/cm}^3$ for an anomaly (also can be a density

³ In geodesy, a coordinate system and a coordinate frame are different in terms of the definitions used (see McCarthy (1996) in Section 3.3). However, in this work, the terms are equally defined and used interchangeably.

⁴ If we have a system of several point masses or solid bodies, the gravitational potential of the system is the sum of the individual contributions (Heiskanen and Moritz, 1979, pp. 1-2).

contrast = an anomaly's density minus a reference density of 2.670 g/cm³), are given as follows:

(a) The point (homogeneous) mass

$$s_{ii}(\mathbf{x}) = \frac{k\rho}{r^3} \frac{(3x_i^2 - r^2)}{r^2} \quad (2.13)$$

and

$$s_{ij}(\mathbf{x}) = \frac{k\rho}{r^3} \frac{3x_i x_j}{r^2} \quad \text{with cyclic indices } (i, j, k) \quad (2.14)$$

(b) The rectangular (homogeneous) prism with edges parallel to the local coordinate axes

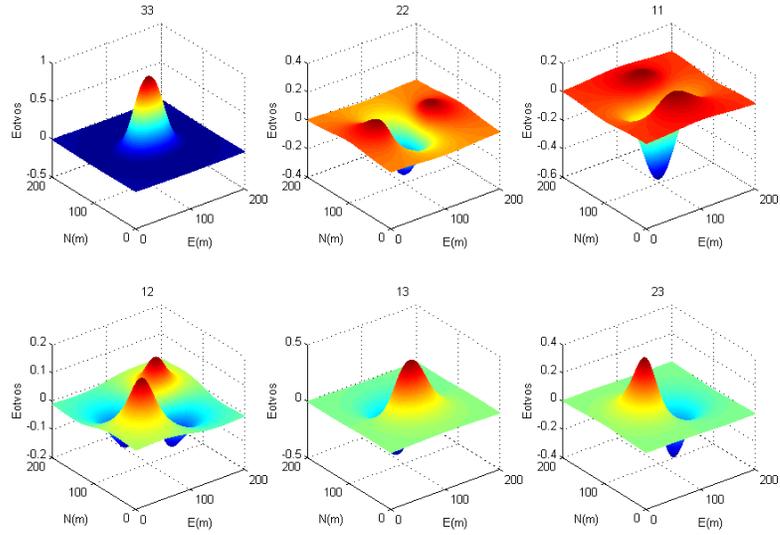
$$s_{ii}(\mathbf{x}) = k\rho \tan^{-1} \left[\frac{(x_j - x'_j)(x_k - x'_k)}{(x_i - x'_i)r} \right] \Big|_{x'_1} \Big|_{x'_2} \Big|_{x'_3} \quad (2.15)$$

and

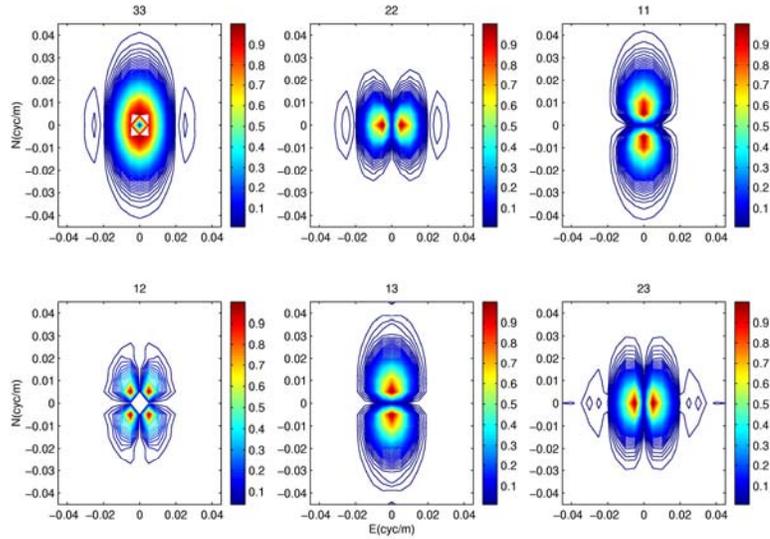
$$s_{ij}(\mathbf{x}) = -k\rho \ln[(x_k - x'_k) + r] \Big|_{x'_1} \Big|_{x'_2} \Big|_{x'_3} \quad (2.16)$$

Since we are interested in the rectangular prism, Figure 2.1a shows only six components of the gravitational gradients due to a rectangular block having a size of 2m x 2m x 50m with density contrast 1.5 g/cm³, computed at 30m above the center of the block at every 2m space interval; the monopole has similar characteristics (not shown here). It is evident that, of all the defined gradients, the s_{33} component has the highest strength of gradient signal; $\max|s_{33}| = 0.91\text{E}$, $\max|s_{13}| = 0.45\text{E}$, $\max|s_{23}| = 0.39\text{E}$, $\max|s_{12}| = 0.14\text{E}$, $\max|s_{11}| = 0.13\text{E}$, and $\max|s_{22}| = 0.11\text{E}$. It should be observed that each component has its own unique characteristics and in turn reflects the shape of the body. In this way, we may recognize what the target anomaly in question looks like. For example, s_{33} has a symmetric and positive envelope. The gradient signals of the s_{11} and s_{22} components have the locations of local maxima appearing at N-S and E-W edges of the prism aligned with the local frame of measurements, respectively. The s_{13} and s_{23} components have two (positive and negative) side lobes along E-W and N-S directions, respectively. On the other hand, s_{12} delineates the corner of the rectangular body - two diagonal pairs of positive spikes and the other two of negative spikes (quadrupole).

It is interesting to look at the individual gradient components due to the rectangular prism in the frequency domain to better understand the signal structure of the body. Figure 2.1b displays the (normalized) absolute amplitudes of the gradients in the frequency domain. The response of s_{33} dominates over a wide range of wavelengths lying between 25m and 1000m along the N-S direction and between 50m and 500m in the E-W direction.



(a)



(b)

Figure 2.1: (a) Six components of gradient tensor due to the 2m x 2m x 50m rectangular box with density contrast 1.5 g/cm^3 , calculated at 30m above the center of the box at every 2m space interval along east direction; (b) their (normalized) absolute amplitudes.

For s_{11} and s_{13} , two main lobes appear along N-S direction. The wavelengths of both span almost the same range as s_{33} . The (magnitude) response of s_{12} has four side lobes (call this a quadrupole). We shall see in Chapter 5 that these particular characteristics of each gradient component will affect the signal detection by using matched filters.

2.4 Transformation

In fact, the feature of the mass anomaly to be detected can exist at any location relative to a local coordinate system, call this the a -frame. For example, consider a main water pipe which runs in a N-E direction (45 degrees with respect to North direction). The gravitational gradients generated by the pipe are equivalent to the gradients due to the pipe parallel to North direction but it is rotated by 45 degrees with respect to the north. The orthogonal transformation of a 3-by-3 matrix representing the gradient tensor is given by (Jekeli, 2000, p. 10-13):

$$s^b = \mathbf{R}_a^b s^a (\mathbf{R}_a^b)^T = \mathbf{R}_a^b s^a \mathbf{R}_b^a \quad (2.17)$$

where s^a and s^b are the gradient tensors in the a -frame and b -frame respectively. \mathbf{R}_a^b represents the rotation (orthogonal) matrix from a -frame to b -frame. The rotation matrix is described by a sequence of rotation for each of the triad axes. The oriented angles about the axes are called Euler angles, consisting of γ about the 1-axis, β about the 2-axis, and α about the 3-axis (or the vertical axis). Positive rotations are considered counter-clockwise by convention and are viewed from the end point of the coordinate frame to the origin.

Rotating the measured gradients at various orientations relative to the survey area yields a preliminary means of getting insight into the target mass anomalies. The rotation matrix works as a filtering tool that directionally separates anomalies and in turn, indicates the anomalies' features (Dransfield, 1994)). For instance, the line feature of the water pipe is clearly identified, among other features, by the rotation. It can be graphically explained in the frequency domain that the rotation separates the particular frequency trend associated with the direction of the pipe, see Figure 5.30 in Chapter 5, where the 2m x 2m x 1000m anomaly is considered in a noise environment.

Before closing this chapter, we want to emphasize that our objectives are to use matched filters to detect a small anomalous mass in strong noise backgrounds. Consider the gravitational gradient signals due to small mass anomalies, e.g., man-made structures and voids, which are very small, compared with the Earth's gradient field. The desired signals can be overwhelmed by unknown mass distributions if the anomalies (i) are not shallow enough and/or (ii) do not have gradient signatures of significant magnitude and uniqueness. What ranges and sizes of such causative sources, sampling intervals (relative to speeds of airplane or helicopter and sampling rate), and flight altitudes (or depths of the sources) impact upon the gradiometer's ability to sense their corresponding signals? These factors are important to conduct an airborne gradiometric survey.

Because the gradient signals of small anomalies are rapidly attenuated with respect to flight altitudes due to the inverse term of high order of range, see eq. (2.13), a gradiometer may not be able to sense the gravitational signals at great depth. The

remainder of dissertation will speculate how these factors can affect the signal detection using matched filters. Characterizing small anomalies relies on the sampling interval; we may lose meaningful signals if the interval is too large, i.e., we introduce aliasing effects. We can envisage that the detection of small anomalies whose gradient signals are a few Eötvös requires not only a very precise gradiometer but also tremendous efforts in statistical signal processing, which will be discussed in the following chapters.

CHAPTER 3

THE THEORY OF SIGNAL DETECTION BY USING MATCHED FILTERS

3.1 Introduction

In this chapter, we shall leave the theories of physical geodesy and geophysics for a moment. Because the detection theory in the field of communication theory may not be familiar for geodesists and geophysicists, we confine our discussion to the concepts of detection theory but applied to our problem at hand. We consider one anomaly embedded in strong geological noise. We discuss only simple hypotheses which describe either the presence or the absence of the signal due to the anomaly.

In this chapter, the ideas of signal detection and matched filters in white and non-white Gaussian noise environments are discussed. Probability related to a likelihood ratio test is discussed as well as hypothesis testing under the Neyman-Pearson criteria. All necessary notations are appropriately defined and used in the entire study unless otherwise specified. For simplicity of derivations, we assume an infinite sequence of measurements unless otherwise specified. In this chapter, an introduction to detection theory relies heavily on some excellent publications on the subject such as by Turin (1960 and 1976), Middleton (1960 and 1965), DiFranco (1968), Van Trees (1968), Whalen (1971), Cadzow (1987), Poor (1983), Helstrom (1960 and 1995), and, Kay (1993 and 1998).

3.2 Signal detection

The concept of signal detection using matched filters for small mass anomaly detection is that it ideally aims to aid in the decision if there exists a wanted gradient signal, s , due to a small anomalous mass at a location \bar{x} (which is unknown). We assume that s is entirely embedded in identically-distributed (i.d.) (or wide sense stationary) zero-mean noise n (not necessarily Gaussian). The two hypotheses are defined as:

(i) Case 0: there is a specific gradient signal in measurements $v(x)$:

$$H_0 : v(x) = s(x - \bar{x}) + n(x) \quad (3.1)$$

(ii) Case 1: there is no gradient signal in measurements $v(x)$:

$$H_1 : v(x) = n(x) \quad (3.2)$$

3.3 Definition of a matched filter

Here, the matched filter (MF) function h of a gradient component is designed such that it locates the signal by determining the maximum in the signal-to-noise ratio (SNR). The value of SNR plays an important role as indicator of the potential location of the signal. For instance, if the SNR is high, it implies the possible existence of the signal. In other words, the MF attempts to enhance s while suppressing n . The optimal weighting coefficients of the MF can be determined using Parseval's theorem and Schwarz's inequality (Papoulis, 1984) as shown below.

In this chapter, we demonstrate a simple type of matched filter when white noise with variance σ^2 is considered. Suppose we have discrete observation $v_j = v(x_j)$, $-\infty < j < +\infty$. Then the discrete filter is formulated as

$$\begin{aligned} y_h(x) &= \sum_{j=-\infty}^{+\infty} h(x-x_j)v_j \\ &= \sum_{j=-\infty}^{+\infty} h(x-x_j)s(x_j-\bar{x}) + \sum_{j=-\infty}^{+\infty} h(x-x_j)n_j \end{aligned} \quad (3.3)$$

The output, y_h , of the filter therefore has a signal part and a noise part. We determine h such that SNR is maximized at \bar{x} . Using (3.3), we define the output SNR as

$$\text{SNR}(\bar{x}) = \frac{\left[\sum_{j=-\infty}^{+\infty} h(\bar{x}-x_j)s(x_j-\bar{x}) \right]^2}{\varepsilon \left\{ \left[\sum_{j=-\infty}^{+\infty} h(\bar{x}-x_j)n_j \right]^2 \right\}} \quad (3.4)$$

where s is a known signal and ε denotes (ensemble) expectation. We further assume that s is square integrable, which implies a finite energy. We make use of the following Fourier Transform pair for infinite discrete signals:

$$\tilde{S}(f) = \Delta x \sum_{j=-\infty}^{+\infty} s_j e^{-i2\pi\Delta x j f} \quad \text{with } -f_N \leq f \leq +f_N \quad (3.5)$$

and

$$s_j = \int_{-f_N}^{+f_N} \tilde{S}(f) e^{i2\pi\Delta x j f} df \quad (3.6)$$

where tilde “ \sim ” indicates the periodicity of spectrum, the Nyquist frequency $f_N = 1/2\Delta x$, and “ Δx ” denotes data spacing. By Parseval's theorem, the numerator of (3.4) can be expressed as

$$\left[\sum_{j=-\infty}^{+\infty} h(\bar{x} - x_j) s(x_j - \bar{x}) \right]^2 = \left[\frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{H}^*(f) e^{-i2\pi\bar{x}f} \tilde{S}^*(f) e^{i2\pi\bar{x}f} df \right]^2 \quad (3.7)$$

For the denominator with, for instance, white noise (i.e., $\varepsilon\{n_j n_i\} = \sigma^2 \delta(i - j)$), where the delta function $\delta(j - i) = 1$ if $j = i$ and $\delta(j - i) = 0$ for $j \neq i$),

$$\begin{aligned} \varepsilon \left\{ \left[\sum_{j=-\infty}^{+\infty} h(\bar{x} - x_j) n_j \right]^2 \right\} &= \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} h(\bar{x} - x_j) h(\bar{x} - x_k) \varepsilon\{n_j n_k\} \\ &= \sigma^2 \sum_{j=-\infty}^{+\infty} [h(\bar{x} - x_j)]^2 \\ &= \frac{\sigma^2}{\Delta x} \int_{-f_N}^{+f_N} |\tilde{H}(f)|^2 df \end{aligned} \quad (3.8)$$

Therefore, with (3.4), (3.7) and (3.8), we apply Schwarz's inequality,

$$\begin{aligned} \text{SNR}(\bar{x}) &= \frac{\left[\frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{H}^*(f) \tilde{S}^*(f) df \right]^2}{\frac{\sigma^2}{\Delta x} \int_{-f_N}^{+f_N} |\tilde{H}(f)|^2 df} \\ &\leq \frac{\frac{1}{\Delta x^2} \int_{-f_N}^{+f_N} |\tilde{H}(f)|^2 df \int_{-f_N}^{+f_N} |\tilde{S}(f)|^2 df}{\frac{\sigma^2}{\Delta x} \int_{-f_N}^{+f_N} |\tilde{H}(f)|^2 df} = \frac{E_s}{\Delta x \sigma^2} \end{aligned} \quad (3.9)$$

where the signal energy

$$E_s = \int_{-f_N}^{+f_N} |\tilde{S}(f)|^2 df \quad (3.10)$$

Clearly, $E_s / \Delta x \sigma^2$ is an upper bound for $\text{SNR}(\bar{x})$ for any h . If we want to maximize $\text{SNR}(\bar{x})$ with respect to h , we find that by choosing

$$\tilde{H}(f) = \frac{1}{\sigma^2} \tilde{S}^*(f) \quad (3.11)$$

we obtain the maximum $\text{SNR}(\bar{x}) = E_s / \Delta x \sigma^2$. The transfer (or response) function of the matched filter (i.e., eq. (3.11)) is the complex conjugate of the spectrum of the signal to which it is matched. For this reason, a matched filter is often called a “conjugate” filter (Turin, 1960). Thus, the filter function h that maximizes $\text{SNR}(\bar{x})$ is given by the inverse Fourier transform of (3.11),

$$h(x_j) = \frac{1}{\sigma^2} s(-x_j) \quad (3.12)$$

or

$$h(x - x_j) = \frac{1}{\sigma^2} s(x_j - x) \quad (3.13)$$

Substituting (3.12) into (3.3), the filter output becomes

$$y_h(x) = \frac{1}{\sigma^2} \sum_{j=-\infty}^{+\infty} s(x_j - x) v_j \quad (3.14)$$

We want to emphasize at this point that the matched filter does not preserve the original shape of the input signal (i.e., s). Since our object is to detect the presence of signal s , we consider the task of extracting the signal at a given point in the background noise. The decision that the prescribed signal may be present can be made wherever the filter output is maximum.

3.4 The matched filter for correlated noise

In this section, we approach a more general form of matched filter. We consider the matched filter for correlated noise, which is derived by a straightforward modification of the derivation of (3.7) and (3.8) noting that eq. (3.4) is also valid for correlated noise. By Wiener-Kinchine theorem, for a wide-sense stationary process, the Fourier transform of the autocorrelation function, $\phi(\cdot)$, is the power spectral density, $\Phi_n(\cdot)$, of the process. Thus, we begin with (3.8) by replacing $\varepsilon\{n_j n_k\}$ with $\phi_{nn}(|x_j - x_k|)$,

$$\begin{aligned}
\varepsilon \left\{ \left[\sum_{j=-\infty}^{+\infty} h(\bar{x} - x_j) n_j \right]^2 \right\} &= \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} h(\bar{x} - x_j) h(\bar{x} - x_k) \phi_{mn}(|x_j - x_k|) \\
&= \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} h(\bar{x} - x_j) h(\bar{x} - x_k) \int_{-f_N}^{+f_N} \Phi_m(f) e^{i2\pi f \Delta x (j-k)} df \\
&= \frac{1}{\Delta x^2} \int_{-f_N}^{+f_N} \Phi_m(f) \left(\left(\Delta x \sum_{k=-\infty}^{+\infty} h(\bar{x} - x_j) e^{-i2\pi f \Delta x (\bar{x} - x_j)} \right) \times \right. \\
&\quad \left. \left(\Delta x \sum_{k=-\infty}^{+\infty} h(\bar{x} - x_j) e^{-i2\pi (-f) \Delta x (\bar{x} - x_k)} \right) \right) df \\
&= \frac{1}{\Delta x^2} \int_{-f_N}^{+f_N} \Phi_m(f) |\tilde{H}(f)|^2 df
\end{aligned} \tag{3.15}$$

Substituting (3.15) into (3.4) and applying Schwarz's inequality,

$$\begin{aligned}
\text{SNR}(\bar{x}) &= \frac{\left[\frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{H}^*(f) \tilde{S}^*(f) df \right]^2}{\frac{1}{\Delta x^2} \int_{-f_N}^{+f_N} \Phi_m(f) |\tilde{H}(f)|^2 df} \\
&\leq \frac{\int_{-f_N}^{+f_N} |\tilde{H}(f)|^2 df \int_{-f_N}^{+f_N} |\tilde{S}(f)|^2 df}{\int_{-f_N}^{+f_N} \Phi_m(f) |\tilde{H}(f)|^2 df}
\end{aligned} \tag{3.16}$$

Clearly, by choosing

$$\tilde{H}(f) = \frac{\tilde{S}^*(f)}{\Phi_m(f)}, \tag{3.17}$$

the left-hand side of the inequality above equals the upper bound on the right-hand side. Thus, we obtain the maximum SNR as follows

$$\text{SNR}_{\max} = \int_{-f_N}^{+f_N} \frac{|\tilde{S}(f)|^2}{\Phi_m(f)} df \tag{3.18}$$

In addition, the expression (3.17) is easily extended to two dimensions in the frequency domain, which will be used for our study, see eq. (5.6) in Chapter 5.

At this point, hypothesis testing is required to evaluate the performance of the matched filter in a statistical sense. This brings us to the concept of classical hypothesis testing which is based on a likelihood ratio test, which requires a priori knowledge of probability density function about the observation (or noise).

3.5 Hypothesis testing under Neyman and Pearson criterion

The Neyman and Pearson theory of hypothesis testing antedates the development of statistical decision theory (DiFranco, 1968, p. 272). It does not require knowledge of a priori signal statistics, nor does it require an explicit assignment of cost functions like Bayes' criterion. Neyman and Pearson define an optimum test as one that minimizes the probability of certain errors. For instance, in a test of hypothesis H , two types of errors can be made: H may be rejected where it is true (i.e., the error is called a type I error), or it may be accepted when it is false (i.e., the error is called a type II error). An optimal test should minimize the probability of both types of errors--the test should have a small probability of rejecting H when it is true and a large probability of rejecting H when it is false. Theoretically, the Neyman and Pearson criterion¹ states that the best test is one that has either the greatest probability of rejecting H when it is false or the greatest probability of accepting H when it is true.

The Neyman and Pearson test is a test between two alternatives hypotheses. For our study, the two hypotheses are (3.1) and (3.2). To further discuss the concept of Neyman and Pearson criteria, let us assume the observations are in some finite interval, $-J \leq j \leq +J$. We start with a likelihood ratio test (LRT) as follows:

$$\text{choose } H_0 \text{ if } \frac{p(\mathbf{v}/H_0)}{p(\mathbf{v}/H_1)} \geq \eta \quad (3.19)$$

versus

$$\text{choose } H_1 \text{ if } \frac{p(\mathbf{v}/H_0)}{p(\mathbf{v}/H_1)} < \eta \quad (3.20)$$

¹ **The Neyman and Pearson criterion**

To minimize POF for a given POM = α'

$$\text{choose } H_0 \text{ if } \frac{p(l/H_0)}{p(l/H_1)} \geq \eta' \quad (a)$$

or

$$\text{choose } H_1 \text{ if } \frac{p(l/H_0)}{p(l/H_1)} < \eta' \quad (b)$$

where η' is obtained from a given α' under the constraint

$$\text{POM} = \int_{-\infty}^{\eta'} p(l/H_0) dl = \alpha' \quad (c)$$

The proof can be found in Kay (1998, pp. 89) and Van Trees (1967, pp. 33-34).

where \mathbf{v} is a vector of measurement. If noise is assumed to be independent and identically distributed, the joint probability density function (pdf) of observations (or the likelihood functions) is just the product of individual probability density functions. Thus, $p(\mathbf{v}/H_1)$ and $p(\mathbf{v}/H_0)$ are called the (joint) conditional pdf's of \mathbf{v} , given H_1 and H_0 respectively. The threshold η is found from a given probability of type I error (see next section).

For simplicity, let n be a white Gaussian random variable with variance σ^2 , then so is \mathbf{v} . Thus, we simply write $p(\mathbf{v}/H_1)$ as the (joint) probability density function of (i.i.d.) noise, given by

$$p(\mathbf{v}/H_1) = p_n(\mathbf{n}) = p_n(n_{-J})p_n(n_{-J+1})p_n(n_{-J+2})\dots p_n(n_j)\dots p_n(n_{+J}) \quad (3.21)$$

where the probability density function of the noise is

$$p_n(n_j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2}\left(\frac{n_j}{\sigma}\right)^2\right\} \quad (3.22)$$

With (3.21) and (3.22), then we have

$$p(\mathbf{v}/H_1) = \frac{1}{(2\pi\sigma^2)^{\frac{2J+1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{j=-J}^{+J} (n_j)^2\right\} \quad (3.23)$$

Similarly, under hypothesis H_0 , where we define $\mathbf{v} = \mathbf{s} + \mathbf{n}$ and then $\mathbf{n} = \mathbf{v} - \mathbf{s}$, with (3.22), we may write

$$p(\mathbf{v}/H_0) = p_n(\mathbf{n}) = p_n(\mathbf{v} - \mathbf{s}) = \frac{1}{(2\pi\sigma^2)^{\frac{2J+1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{j=-J}^{+J} (v_j - s_j)^2\right\} \quad (3.24)$$

Using (3.1), (3.2), (3.23) and (3.24), the corresponding likelihood ratio is

$$\frac{p(\mathbf{v}/H_0)}{p(\mathbf{v}/H_1)} = \exp\left\{-\frac{1}{2\sigma^2} \sum_{j=-J}^{+J} [(v(x_j) - s(x_j - \bar{x}))^2 - (v(x_j))^2]\right\} \quad (3.25)$$

Since the exponential function in (3.25) is monotonic, taking the natural logarithm does not change the inequalities of (3.19) and (3.20):

$$\ln\left[\frac{p(\mathbf{v}/H_0)}{p(\mathbf{v}/H_1)}\right] = \frac{1}{\sigma^2} \sum_{j=-J}^{+J} v(x_j)s(x_j - \bar{x}) - \frac{1}{2\sigma^2} \sum_{j=-J}^{+J} s(x_j - \bar{x})^2 \quad (3.26)$$

The second term on the right hand side of (3.26) is just the energy of the signal. Only the first term significantly describes the probabilistic nature of the observations. This term is called a *sufficient statistic*, say l , (Van Trees, 1968, p. 29). The hypothesis testing criteria of (3.19) and (3.20) become:

$$\text{choose } H_0 \text{ if } l(\bar{x}) \geq \eta' \quad (3.27)$$

versus

$$\text{choose } H_1 \text{ if } l(\bar{x}) < \eta' \quad (3.28)$$

where

$$l(\bar{x}) = \frac{1}{\sigma^2} \sum_{j=-J}^{+J} v(x_j) s(x_j - \bar{x}) \quad (3.29)$$

and

$$\eta' = \ln \eta + \frac{1}{2\sigma^2} \sum_{j=-J}^{+J} s(x_j - \bar{x})^2 \quad (3.30)$$

By comparing (3.14) with (3.29), we immediately see that the sufficient statistic, for the case of white Gaussian noise, corresponds to the output of the matched filter at the location of the signal. According to the Neyman-Pearson criterion, either η or η' can be assigned or is dependent on the predetermined value of a significant level, say α , that will be elaborated in the next section.

3.6 Performance of detection

As mentioned in the previous section, signal detection using the matched filter can be assessed statistically using hypothesis testing through the sufficient statistic of (3.29). Figure 3.1 describes *error probability densities*, which lead to very important terms in detection theory as follows. We define the probability of *miss*, POM, or the probability of a Type I error in this case, as corresponding to rejecting H_0 when it is true. We also define the probability of *false alarm*, POF, or the probability of a Type II error in this case, which refers to choosing H_0 when there is no signal. The probability of *detection*, POD, corresponds to the correct decision and relates to POM as

$$\text{POD} = 1 - \text{POM} \quad (3.31)$$

The implicit forms of the error probabilities are given by (see next sections for more details):

$$\text{POF} = \int_{R_0} p(l/H_1) dv, \quad (3.32)$$

$$\text{POD} = \int_{R_0} p(l/H_0) dv, \quad (3.33)$$

and

$$\text{POM} = \int_{R_1} p(l/H_0) dv. \quad (3.34)$$

where p denotes a probability density and R_0 and R_1 denote the decision regions of choosing H_0 and H_1 , respectively, and the integral notations are valid for either discrete or continuous points. We can connect POF, POD, and POM with a sufficient statistic as in the following sections.

3.6.1 White noise case

For a large amount of data, we may assume $-\infty < j < +\infty$. Thus, the equations of (3.29) and (3.30) for an arbitrary point x , can be rewritten as follows:

$$l(x) = \frac{1}{\sigma^2} \sum_{j=-\infty}^{+\infty} v(x_j) s(x_j - x) \quad (3.35)$$

and

$$\eta' = \ln \eta + \frac{1}{2\sigma^2} \sum_{j=-\infty}^{+\infty} s(x_j - x)^2 \quad (3.36)$$

where η' is obtained from a given POM (see (3.43)). Because $l(x)$ is a linear operation on $v(x)$, it is still a (white) Gaussian random variable. Following Parseval's theorem, it is straightforward to show that, the *conditional* expectation of l on H_0 corresponds to SNR_{\max} in (3.9):

$$\begin{aligned} \varepsilon\{l/H_0\} &= \varepsilon\left\{\frac{1}{\sigma^2} \sum_{j=-\infty}^{+\infty} s(x_j - x)v(x_j) \middle/ H_0\right\} \\ &= \frac{1}{\sigma^2} \sum_{j=-\infty}^{+\infty} s(x_j - x)\varepsilon\{v(x_j)/H_0\} \\ &= \frac{1}{\sigma^2} \sum_{j=-\infty}^{+\infty} s(x_j - x)s(x_j - x) \\ &= \frac{E_s}{\Delta x \sigma^2} \\ &= d^2 \end{aligned} \quad (3.37)$$

and, $\varepsilon\{l/H_1\} = 0$. The *conditional* covariance (variance) of l under either hypothesis is the same. We have

$$\begin{aligned}
Var\{l/H_{0,1}\} &= Var\left\{\frac{1}{\sigma^2} \sum_{j=-\infty}^{+\infty} s(x_j - x)v(x_j) / H_{0,1}\right\} \\
&= \frac{1}{\sigma^4} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} s(x_j - x)s(x_k - x)C\{v(x_j)v(x_k)/H_{0,1}\} \\
&= \frac{1}{\sigma^4} \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} s(x_j - x)s(x_k - x)\sigma^2\delta(j - k) \\
&= \frac{1}{\sigma^2} \sum_{j=-\infty}^{+\infty} s(x_j - x)s(x_j - x) \\
&= \frac{E_s}{\Delta x \sigma^2} \\
&= d^2
\end{aligned} \tag{3.38}$$

Finally, as shown in Figure 3.1, we select Type I error (i.e., $\alpha = \text{POM}$) to compute a threshold. We have

$$\begin{aligned}
\text{POD} &= \int_{R_0'} p(l/H_0) dl \\
&= \int_{\eta'}^{+\infty} \frac{1}{\sqrt{2\pi}d^2} \exp\left[-\frac{1}{2} \frac{(l-d^2)^2}{d^2}\right] dl \\
&= \int_{\eta'}^{+\infty} \frac{1}{\sqrt{2\pi}d^2} \exp\left[-\frac{1}{2} \left(\frac{l}{d} - d\right)^2\right] dl
\end{aligned} \tag{3.39}$$

Then,

$$\text{POM} = 1 - \int_{\eta'}^{+\infty} \frac{1}{\sqrt{2\pi}d^2} \exp\left[-\frac{1}{2} \left(\frac{l}{d} - d\right)^2\right] dl \tag{3.40}$$

Equation (3.40) can be rewritten in terms of the complement to the error function, *erfc*,

$$\text{POM} = 1 - 0.5 \text{erfc}\left[\frac{1}{\sqrt{2}} \left(\frac{\eta'}{d} - d\right)\right], \tag{3.41}$$

where

$$\text{erfc}(Z) = \int_Z^{+\infty} \frac{2}{\sqrt{\pi}} \exp[-x^2] dx \tag{3.42}$$

By selecting POM, the threshold can be computed as

$$\eta' = d^2 + \sqrt{2d^2} \operatorname{erfc}^{-1}[2 - 2\text{POM}], \quad (3.43)$$

Then, we can compute Type II error ($\beta = \text{POF}$) from (3.43)

$$\text{POF} = 0.5 \operatorname{erfc} \left[\frac{\eta'}{\sqrt{2d^2}} \right] \quad (3.44)$$

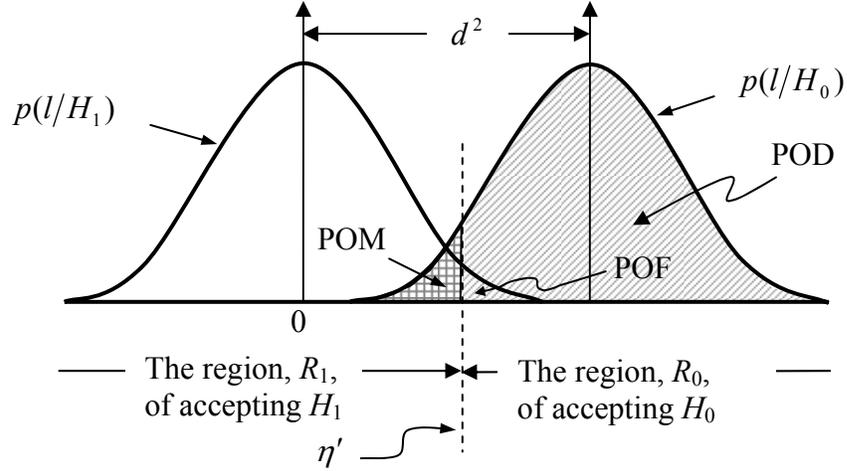


Figure 3.1: Error probabilities in Gaussian distribution: POF, POD, and POM.

3.6.2 Correlated noise case

The derivation of detection performance in stationary-non-white Gaussian noise is more complicated. Let \mathbf{v} be a $N \times 1$ vector of the measurement consisting of the $N \times 1$ signal vector \mathbf{s} and the $N \times 1$ (additive) noise vector \mathbf{n} . To simplify the derivation, we assume that the covariance matrix $\mathbf{\Sigma}$ of the noise is positive definite. Then, the loglikelihood ratio is given by

$$\ln \left[\frac{p(\mathbf{v}/H_0)}{p(\mathbf{v}/H_1)} \right] = \ln \left[\frac{\frac{1}{(2\pi)^N |\mathbf{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{v} - \mathbf{s})^T \mathbf{\Sigma}^{-1} (\mathbf{v} - \mathbf{s}) \right]}{\frac{1}{(2\pi)^N |\mathbf{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2} \mathbf{v}^T \mathbf{\Sigma}^{-1} \mathbf{v} \right]} \right] \quad (3.45)$$

Simplifying:

$$\begin{aligned} \ln \left[\frac{p(\mathbf{v}/H_0)}{p(\mathbf{v}/H_1)} \right] &= -\frac{1}{2} \left[(\mathbf{v} - \mathbf{s})^T \boldsymbol{\Sigma}^{-1} (\mathbf{v} - \mathbf{s}) - \mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{v} \right] \\ &= \mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} - \frac{1}{2} \mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \end{aligned} \quad (3.46)$$

Therefore, the first term of (3.46) is the sufficient statistic l . Then, decide H_0 if

$$l = \mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} > \eta' = \ln \eta + \frac{1}{2} \mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \quad (3.47)$$

The conditional expectations of the sufficient statistic l are

$$\varepsilon\{l/H_1\} = \varepsilon\{\mathbf{n}^T \boldsymbol{\Sigma}^{-1} \mathbf{s}\} = \mathbf{s}^T \boldsymbol{\Sigma}^{-1} \varepsilon\{\mathbf{n}\} = 0 \quad (3.48)$$

and

$$\varepsilon\{l/H_0\} = \varepsilon\{(\mathbf{s} + \mathbf{n})^T \boldsymbol{\Sigma}^{-1} \mathbf{s}\} = \mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \quad (3.49)$$

The conditional variances are

$$\text{Var}\{l/H_1\} = \varepsilon\{(\mathbf{n}^T \boldsymbol{\Sigma}^{-1} \mathbf{s})^2\} = \mathbf{s}^T \boldsymbol{\Sigma}^{-1} \varepsilon\{\mathbf{n} \mathbf{n}^T\} \boldsymbol{\Sigma}^{-1} \mathbf{s} = \mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \quad (3.50)$$

and

$$\begin{aligned} \text{Var}\{l/H_0\} &= \varepsilon\left\{ \left(\mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} - \varepsilon\{\mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{s}\} \right)^2 \right\} \\ &= \varepsilon\left\{ \left(\mathbf{v}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} - \varepsilon\{\mathbf{v}^T\} \boldsymbol{\Sigma}^{-1} \mathbf{s} \right)^2 \right\} \\ &= \varepsilon\left\{ \left((\mathbf{v} - \varepsilon\{\mathbf{v}\})^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \right)^2 \right\} \\ &= \varepsilon\left\{ \left((\mathbf{v} - \mathbf{s})^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \right)^2 \right\} \\ &= \varepsilon\left\{ \left(\mathbf{n}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \right)^2 \right\} \\ &= \text{Var}\{l/H_1\} \end{aligned} \quad (3.51)$$

As in the white noise case, we can easily find POF from a given POM. In practice, it should be noted that it is not possible to invert the covariance matrix $\boldsymbol{\Sigma}$ in (3.47) and also (3.49) if the data vector is large. In fact, Fourier transforms of \mathbf{v} and \mathbf{n} are formed by linear operations on the Gaussian random variables \mathbf{v} and \mathbf{n} respectively. The Fourier transforms themselves are also Gaussian (complex) variables. By applying the orthogonal transformation matrix, \mathbf{F} , to (3.47) and (3.49) (see section 6.3.1 for further details), the corresponding covariance matrix becomes diagonal--only the diagonal elements of the matrix are inverted:

$$\begin{aligned}
l &= \mathbf{v}^T (\mathbf{F}\mathbf{F}^H \boldsymbol{\Sigma} \mathbf{F}\mathbf{F}^H)^{-1} \mathbf{s} \\
&= \mathbf{s}^T (\mathbf{F}\mathbf{F}^H \boldsymbol{\Sigma} \mathbf{F}\mathbf{F}^H)^{-1} \mathbf{v} \\
&= \mathbf{s}^T \mathbf{F}\mathbf{F}^H \boldsymbol{\Sigma}^{-1} \mathbf{F}\mathbf{F}^H \mathbf{v} \\
&= \frac{\Delta x \Delta x}{\Delta x \Delta x} (\mathbf{F}^H \mathbf{s})^T (\mathbf{F}^H \boldsymbol{\Sigma}^{-1} \mathbf{F})^{-1} (\mathbf{F}^H \mathbf{v}) \\
&= \frac{1}{N \Delta x} \sum_{i=0}^{N-1} \tilde{S}^*(f_i) \Delta x^{-1} \tilde{\Sigma}^{-1}(f_i) \tilde{V}(f_i)
\end{aligned} \tag{3.52}$$

and also

$$\begin{aligned}
\mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} &= (\mathbf{F}^H \mathbf{s})^H (\mathbf{F}^H \boldsymbol{\Sigma} \mathbf{F})^{-1} (\mathbf{F}^H \mathbf{s}) \\
&= \frac{1}{N \Delta x} \sum_{i=0}^{N-1} \tilde{S}^*(f_i) \Delta x^{-1} \tilde{\Sigma}^{-1}(f_i) \tilde{S}(f_i)
\end{aligned} \tag{3.53}$$

where S and V are Fourier transforms of s and v after applying the orthogonal transformation matrix \mathbf{F} , and N is the number of measurement points along track. For a large data vector, we may approximate $\tilde{\Sigma}(f) \approx \Delta x^{-1} \Phi_{mn}(f)$. As such, we have

$$l \approx \int_{-f_N}^{+f_N} \frac{\tilde{V}(f) \tilde{S}^*(f)}{\Phi_{mn}(f)} df \tag{3.54}$$

and

$$\mathbf{s}^T \boldsymbol{\Sigma}^{-1} \mathbf{s} \approx \int_{-f_N}^{+f_N} \frac{|\tilde{S}(f)|^2}{\Phi_{mn}(f)} df = \text{SNR}_{\max} = d^2 \tag{3.55}$$

Thus, POF can be computed using (3.41) to (3.44) with (3.55). The expressions above can be extended to two dimensions in the frequency domain. Let M and N be the number of data tracks and the number of points of measurements along track, respectively. For this case, the symbols \mathbf{v} and \mathbf{s} are the $MN \times 1$ vectors of observation and signal respectively. The $MN \times MN$ cross-covariance matrix of the $MN \times 1$ noise vector \mathbf{n} is $\boldsymbol{\Sigma}'$. Similar to the derivations of (3.52) through (3.55) but applying the orthogonal transformation matrix \mathbf{F}' , described in section 6.3.3 to (3.47) and (3.49) instead, we finally arrive at

$$l' \approx \int_{-f_N}^{+f_N} \int_{-f_N}^{+f_N} \tilde{S}^*(f_i, f_j) \Phi_{mn}^{-1}(f_i, f_j) \tilde{V}(f_i, f_j) df df' \tag{3.56}$$

and

$$\mathbf{s}^T \boldsymbol{\Sigma}'^{-1} \mathbf{s} \approx \int_{-f'_N}^{+f'_N} \int_{-f_N}^{+f_N} \frac{|\tilde{S}(f, f')|^2}{\Phi_m(f, f')} df df' = \text{SNR}_{\max} = d^2 \quad (3.57)$$

where S and V are Fourier transforms of s and v after applying the orthogonal transformation matrix \mathbf{F}' . Then, POF can be easily obtained from (3.41) to (3.44) with (3.57).

CHAPTER 4

THE OPTIMAL COMBINATION OF SIX COMPONENTS OF GRAVITATIONAL GRADIENTS

4.1 Introduction

Thus far, we have learned how the six components of the gravitational gradient tensor describe the local structure of gradient field. Each component represents particular characteristics at the same location where measurements are obtained by the gradiometer. With multiple sensors, one would expect improved detection of gradient signals in a noisy background. However, independent filters in each gradient seem to be impractical for the case of detecting small anomalies in a complex gradient field (numerical analyses are discussed in Chapter 5). Therefore, the idea of a vector random process is considered to enhance the capabilities of matched filtering. We consider the matched filter for a combination of six gradient components, with full specification covariances between the individual gradients. The approach processes the combination data in the frequency domain using Fourier transforms. With the property of uncorrelated coefficients at different frequencies, the block diagonal matrices to be inverted have the size of 6x6 at most, corresponding to the 6 gradient tensor elements. For simplicity in the derivations, we assume an infinite sequence of measurements. In practice, however, the analysis is performed on a finite amount of data, which are further assumed to be periodic. Therefore, the matched filter solution may not be optimal (i.e., it is an approximation to the solution).

4.2 The combination of matched filter

We can generalize the matched filters, discussed in Chapter 3, to the multi-sensor case. From (3.1), we extend the measurement v_p for sensor p , with $p = \{1, 2, \dots, P\}$, containing the infinite sequence signal s_p and noise n_p . The observation becomes

$$v_p(x_j) = s_p(x_j - \bar{x}) + n_p(x_j) \quad \text{with } -\infty < j < +\infty \quad (4.1)$$

or in a vector format

$$\mathbf{v}(x_j) = \mathbf{s}(x_j - \bar{x}) + \mathbf{n}(x_j) \quad (4.2)$$

where each vector has the size of $P \times 1$ at point x_j .

Here, we desire to construct a filter that in its output indicates potential locations \bar{x} for the signal. The filter has the form

$$\mathbf{y}(x) = \sum_{j=-\infty}^{+\infty} \mathbf{h}(x - x_j) \mathbf{v}(x_j) \quad (4.3)$$

where now \mathbf{h} is a $P \times P$ matrix function and \mathbf{y} is a $P \times 1$ vector function. As before, we define signal and noise components of the output

$$\mathbf{y}(x) = \sum_{j=-\infty}^{+\infty} \mathbf{h}(x - x_j) s(x_j - \bar{x}) + \sum_{j=-\infty}^{+\infty} \mathbf{h}(x - x_j) \mathbf{n}(x_j) \quad (4.4)$$

and we need to define a SNR which at a particular $x = \bar{x}$ is maximized, thus determining \mathbf{h} . We define the SNR as follows

$$\text{SNR}(\bar{x}) = \frac{\left[\sum_{j=-\infty}^{+\infty} \mathbf{h}(\bar{x} - x_j) s(x_j - \bar{x}) \right]^T \left[\sum_{j=-\infty}^{+\infty} \mathbf{h}(\bar{x} - x_j) s(x_j - \bar{x}) \right]}{\mathcal{E} \left\{ \left[\sum_{j=-\infty}^{+\infty} \mathbf{h}(\bar{x} - x_j) \mathbf{n}(x_j) \right]^T \left[\sum_{j=-\infty}^{+\infty} \mathbf{h}(\bar{x} - x_j) \mathbf{n}(x_j) \right] \right\}} \quad (4.5)$$

Note that numerator and denominator are vector dot-products, i.e., scalars. In the numerator, the components of each vector are

$$\sum_{j=-\infty}^{+\infty} \mathbf{h}_p^T(\bar{x} - x_j) \mathbf{n}(x_j) \quad (4.6)$$

where \mathbf{h}_p^T is the p^{th} vector of \mathbf{h} . By Parseval's theorem, we then have

$$\begin{aligned} \sum_{j=-\infty}^{+\infty} \mathbf{h}_p^T(\bar{x} - x_j) s(x_j - \bar{x}) &= \sum_{q=1}^P \sum_{j=-\infty}^{+\infty} h_{pq}(\bar{x} - x_j) s_q(x_j - \bar{x}) \\ &= \sum_{q=1}^P \frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{H}_{pq}^*(f) \tilde{S}_q^*(f) df \\ &= \frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\mathbf{S}}^*(f) df \end{aligned} \quad (4.7)$$

where $\tilde{\mathbf{H}}_p^*(f) = [\tilde{H}_{p1}^*(f) \dots \tilde{H}_{pp}^*(f)]^T$ and $\tilde{\mathbf{S}}^*(f) = [\tilde{S}_1^*(f) \dots \tilde{S}_p^*(f)]^T$. This is the p^{th} component of the vector appearing in the numerator of the SNR. Therefore, the numerator is given by

$$\frac{1}{\Delta x^2} \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_1^*(f)^T \tilde{\mathbf{S}}^*(f) df \right)^2 + \dots + \frac{1}{\Delta x^2} \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\mathbf{S}}^*(f) df \right)^2 \quad (4.8)$$

For the denominator of the SNR, we proceed similarly. Again, it is the expectation of a dot-product of vectors, whose components are

$$\sum_{j=-\infty}^{+\infty} \mathbf{h}_p^T(\bar{x} - x_j) \mathbf{n}(x_j) = \sum_{q=1}^P \sum_{j=-\infty}^{+\infty} h_{pq}(\bar{x} - x_j) n_q(x_j) \quad (4.9)$$

Therefore, the expectation of the dot-product is given by

$$\mathcal{E} \left\{ \begin{aligned} & \sum_{q=1}^P \sum_{q'=1}^P \sum_{j=-\infty}^{+\infty} \sum_{j'=-\infty}^{+\infty} h_{1q}(\bar{x} - x_j) h_{1q'}(\bar{x} - x_{j'}) n_q(x_j) n_{q'}(x_{j'}) + \dots \\ & \dots + \sum_{q=1}^P \sum_{q'=1}^P \sum_{j=-\infty}^{+\infty} \sum_{j'=-\infty}^{+\infty} h_{Pq}(\bar{x} - x_j) h_{Pq'}(\bar{x} - x_{j'}) n_q(x_j) n_{q'}(x_{j'}) \end{aligned} \right\} \quad (4.10)$$

Taking the expectation inside the sums, according to the assumption of stationary process in Chapter 3, we can write

$$\mathcal{E}\{n_q(x_j) n_{q'}(x_{j'})\} = \phi_{n_q n_{q'}}(|x_j - x_{j'}|) \quad (4.11)$$

with the $P \times P$ covariance matrix

$$\boldsymbol{\Sigma}_{nn} = \begin{bmatrix} \phi_{n_1 n_1} & \dots & \phi_{n_1 n_P} \\ \vdots & \ddots & \vdots \\ \phi_{n_P n_1} & \dots & \phi_{PP} \end{bmatrix} \quad (4.12)$$

We allow correlated noise, i.e., this could be a combination of geologic background plus observation error (white or correlated). For one of the quadruple sums (4.10), we have

$$\begin{aligned} & \sum_{q=1}^P \sum_{q'=1}^P \sum_{j=-\infty}^{+\infty} h_{pq}(\bar{x} - x_j) \left(\sum_{j'=-\infty}^{+\infty} h_{pq'}(\bar{x} - x_{j'}) \phi_{n_q n_{q'}}(|x_j - x_{j'}|) \right) \\ & = \sum_{q=1}^P \sum_{q'=1}^P \sum_{j=-\infty}^{+\infty} h_{pq}(\bar{x} - x_j) \left(\frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{H}_{pq'}(f) \tilde{\Phi}_{n_q n_{q'}}(f) e^{-i2\pi\Delta x f j} e^{i2\pi\bar{x} f} df \right) \end{aligned} \quad (4.13)$$

$$\begin{aligned}
&= \sum_{q=1}^P \sum_{q'=1}^P \frac{1}{\Delta x^2} \int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_{pq}^*(f) \tilde{\Phi}_{n_q n_{q'}}(f) \tilde{\mathbf{H}}_{pq'}(f) df \\
&= \frac{1}{\Delta x^2} \int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\Phi}_m(f) \tilde{\mathbf{H}}_p(f) df
\end{aligned}$$

With (4.13), the denominator of (4.10) is given by

$$\frac{1}{\Delta x^2} \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_1^*(f)^T \tilde{\Phi}_m(f) \tilde{\mathbf{H}}_1(f) df \right)^2 + \dots + \frac{1}{\Delta x^2} \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\Phi}_m(f) \tilde{\mathbf{H}}_p(f) df \right)^2 \quad (4.14)$$

where $\tilde{\Phi}_m$ is the $P \times P$ cross power spectral density matrix.

We revisit the numerator of the SNR and write each sum as

$$\frac{1}{\Delta x^2} \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\mathbf{S}}^*(f) df \right)^2 = \frac{1}{\Delta x^2} \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\Phi}_m^{\frac{1}{2}}(f) \tilde{\Phi}_m^{-\frac{1}{2}}(f) \tilde{\mathbf{S}}^*(f) df \right)^2 \quad (4.15)$$

Again, by Schwarz's inequality, this is

$$\leq \frac{1}{\Delta x^2} \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\Phi}_m^{\frac{1}{2}}(f) \tilde{\Phi}_m^{-\frac{1}{2}}(f) \tilde{\mathbf{H}}_p(f) df \right) \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f)^T \tilde{\Phi}_m^{-\frac{1}{2}}(f) \tilde{\Phi}_m^{\frac{1}{2}}(f) \tilde{\mathbf{S}}(f) df \right) \quad (4.16)$$

Hence, with (4.14), the SNR is bounded by

$$\begin{aligned}
\text{SNR} &\leq \frac{\frac{1}{\Delta x^2} \sum_{p=1}^P \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\Phi}_m(f) \tilde{\mathbf{H}}_p(f) df \right) \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f)^T \tilde{\Phi}_m^{-1}(f) \tilde{\mathbf{S}}(f) df \right)}{\frac{1}{\Delta x^2} \sum_{p=1}^P \left(\int_{-f_N}^{+f_N} \tilde{\mathbf{H}}_p^*(f)^T \tilde{\Phi}_m(f) \tilde{\mathbf{H}}_p(f) df \right)} \\
&\leq \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f)^T \tilde{\Phi}_m^{-1}(f) \tilde{\mathbf{S}}(f) df
\end{aligned} \quad (4.17)$$

As before, if we choose

$$\tilde{\mathbf{H}}_p(f) = \tilde{\Phi}_m^{-1}(f) \tilde{\mathbf{S}}^*(f) \text{ for all } p \quad (4.18)$$

then the numerator of the SNR becomes

$$\left(\frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f)^T \tilde{\Phi}_{nm}^{-1}(f) \tilde{\mathbf{S}}(f) df \right)^2 \quad (4.19)$$

and the denominator of the SNR becomes

$$\left(\frac{1}{\Delta x} \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f)^T \tilde{\Phi}_{nm}^{-1}(f) \tilde{\mathbf{S}}(f) df \right) \quad (4.20)$$

Therefore,

$$\text{SNR} = \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f)^T \tilde{\Phi}_{nm}^{-1}(f) \tilde{\mathbf{S}}(f) df \quad (4.21)$$

which is the maximum value. Thus, for SNR to be the maximum at \bar{x} we choose the transfer function of the combined filter according to (4.18). The filter is the same for all outputs of the vector \mathbf{y} , which means the filter transforms all inputs from multiple sensors into one-dimensional output. We can simply define the combination of matched filters (vector) as

$$\mathbf{y}(x) = \sum_{j=-\infty}^{+\infty} \bar{\mathbf{h}}^T(x - x_j) \mathbf{v}(x_j) \quad (4.22)$$

where

$$\bar{\mathbf{h}}(x) = \int_{-f_N}^{+f_N} \tilde{\Phi}_{nm}^{-1}(f) \tilde{\mathbf{S}}^*(f) e^{i2\pi x f} df \quad (4.23)$$

The derivations discussed above are for the case of a single data track and multiple sensors (SM) (see also Section 6.3.2 in Chapter 6). For multiple data track and multiple sensors (MS), we simply extend (4.22) and the transfer function of (4.23) such that

$$\mathbf{y}(x, x') = \sum_{i=-\infty}^{+\infty} \sum_{j=-\infty}^{+\infty} \bar{\mathbf{h}}^T(x - x_i, x' - x'_j) \mathbf{v}(x_i, x'_j) \quad (4.24)$$

with

$$\bar{\mathbf{h}}(x, x') = \int_{-f_N}^{+f_N} \int_{-f'_N}^{+f'_N} \tilde{\Phi}_{nm}^{-1}(f, f') \tilde{\mathbf{S}}^*(f, f') e^{i2\pi(xf + x'f')} df' df \quad (4.25)$$

where x' is a point across track and Φ_{nm} has the size of $P \times P$.

Following the same procedure as for the single sensor in Chapter 3, the next section will show the connection between the combination of matched filter and hypothesis

testing under the Neyman and Pearson criterion. The derivations in turn indicate that y corresponds to a sufficient statistic in a log-likelihood ratio test. Although it is rather tedious to do so, still, the derivations reveal a significant contribution concerning the inversion of the covariance matrix in the frequency domain (intuitive details, also, can be found in Chapter 6).

4.3 Hypothesis testing and performance of detection

Similar to section 3.7, with (4.2) and (4.12) we assume some finite amount of data for the convenience of derivations, e.g., N data points. Then, the sufficient statistic l_{SM} for a single track and multiple sensors can be given by

$$l_{SM} = \mathbf{v}^T \boldsymbol{\Sigma}_{SM}^{-1} \mathbf{s} \quad (4.26)$$

where $\boldsymbol{\Sigma}_{SM}$ is the $PN \times PN$ covariance matrix of the $PN \times 1$ noise vector \mathbf{n} . Then, the conditional expectations and variances of l_{SM} under H_0 and H_1 can be obtained (see also (3.48) to (3.52) for comparison). Using the orthogonal transformation \mathbf{F}_{SM} as (C.11) in Appendix C the transformed covariance matrix $\mathbf{F}_{SM}^H \boldsymbol{\Sigma}_{SM} \mathbf{F}_{SM}$ becomes diagonal,

$$\begin{aligned} l_{SM} &= \mathbf{v}^T \left(\mathbf{F}_{SM} \mathbf{F}_{SM}^H \boldsymbol{\Sigma}_{SM} \mathbf{F}_{SM} \mathbf{F}_{SM}^H \right)^{-1} \mathbf{s} \\ &= \frac{\Delta x \Delta x}{\Delta x \Delta x} \left(\mathbf{F}_{SM}^H \mathbf{s} \right)^T \left(\mathbf{F}_{SM}^H \boldsymbol{\Sigma}_{SM}^{-1} \mathbf{F}_{SM} \right)^{-1} \left(\mathbf{F}_{SM}^H \mathbf{v} \right) \\ &= \frac{1}{N \Delta x} \sum_{i=0}^{N-1} \tilde{\mathbf{S}}^*(f_i) \Delta x^{-1} \tilde{\boldsymbol{\Sigma}}_{SS}^{-1}(f_i) \tilde{\mathbf{V}}(f_i) \end{aligned} \quad (4.27)$$

For a large amount of data, we may approximate $\mathbf{F}_{SM}^H \boldsymbol{\Sigma}_{SM} \mathbf{F}_{SM} \approx \Delta x^{-1} \tilde{\Phi}_m$ (see also (6.65)). Then,

$$l_{SM} \approx \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f) \tilde{\Phi}_m^{-1}(f) \tilde{\mathbf{V}}(f) df \quad (4.28)$$

and

$$\mathbf{s}^T \boldsymbol{\Sigma}_{SM}^{-1} \mathbf{s} \approx \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f)^T \tilde{\Phi}_m^{-1}(f) \tilde{\mathbf{S}}(f) df = \text{SNR}_{\max} = d^2 \quad (4.29)$$

Similar to the derivations of (4.26) through (4.29), for the case of multiple data tracks, say N' tracks with track spacing $\Delta x'$, and multiple sensors, we have (see also Section 6.3.4 in Chapter 6)

$$\begin{aligned}
l_{\text{MM}} &= \mathbf{v}^T (\mathbf{F}_{\text{MM}} \mathbf{F}_{\text{MM}}^H \boldsymbol{\Sigma}_{\text{MM}} \mathbf{F}_{\text{MM}} \mathbf{F}_{\text{MM}}^H)^{-1} \mathbf{s} \\
&= (\mathbf{F}_{\text{MM}}^H \mathbf{s})^T (\mathbf{F}_{\text{MM}}^H \boldsymbol{\Sigma}_{\text{MM}}^{-1} \mathbf{F}_{\text{MM}})^{-1} (\mathbf{F}_{\text{MM}}^H \mathbf{v}) \\
&= \frac{1}{N \Delta x N' \Delta x'} \sum_{i=0}^{N-1} \sum_{j=0}^{N'-1} \tilde{\mathbf{S}}^*(f_i, f_j) \Delta x^{-1} \Delta x'^{-1} \tilde{\boldsymbol{\Sigma}}_{\text{MM}}^{-1}(f_i, f_j) \tilde{\mathbf{V}}(f_i, f_j)
\end{aligned} \tag{4.30}$$

For a large amount of data,

$$l_{\text{MM}} \approx \int_{-f'_N}^{+f'_N} \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f, f') \tilde{\boldsymbol{\Phi}}_m^{-1}(f, f') \tilde{\mathbf{V}}(f, f') df df' \tag{4.31}$$

and

$$\mathbf{s}^T \boldsymbol{\Sigma}_{\text{MM}}^{-1} \mathbf{s} \approx \int_{-f'_N}^{+f'_N} \int_{-f_N}^{+f_N} \tilde{\mathbf{S}}^*(f, f') \tilde{\boldsymbol{\Phi}}_m^{-1}(f, f') \tilde{\mathbf{S}}(f, f') df df' = \text{SNR}_{\text{max}} = d^2 \tag{4.32}$$

CHAPTER 5

NUMERICAL DISCUSSIONS ON SHALLOW SMALL ANOMALOUS MASS DETECTION

5.1 The simulated survey of airborne gradiometry

Since a precise gradiometer that measures all components of the gravitational gradient tensor, with overall noise performance of about $1 \text{ E}/\sqrt{\text{Hz}}$ is still under development; actual gradiometric data with such high precision do not exist in an operational setting. For this work, six components of the gravitational gradient tensor are simulated by using the 1987 GGSS data set and the gradients computed from a digital elevation model (DEM). Although the GGSS system was corrupted by vehicle vibrations and data outages along several survey tracks, the test data demonstrated an extraordinary accomplishment of technology and associated signal processing (Jekeli, 1986). Vasco (1989) and Vasco and Taylor (1991) showed that the GGSS gradient data reflect subsurface structures as deep as 12km over the Texas/Oklahoma area. The survey area for the simulation is chosen for its moderate terrain surface; see Table 5.1. The area has the size of 36km x 36km and lies above the Meunster uplift. The densities of the region, where the survey was planned, vary in range from 2.500 to 2.960 g/cm³. For simplicity, we use the average density of 2.670 g/cm³. The DEM data sets used in this work are the 1 arcsec National Elevation Dataset (NED), which are produced by the U.S. geological survey¹ (USGS).

Having selected the area site, the gradiometric survey can be designed. The survey data should be sufficiently dense so as to allow the detection of meaningful gradient signals due to small mass anomalies and in turn to avoid aliasing effects. However, the sampling spaces and track spacing should be large enough to represent reasonably cost-effective and timely surveys. We choose the helicopter velocity, v , of 10m/s with the 3-second sampling interval, τ , (in other words, the sampling frequency, $\Delta_f = 1/\tau$, is 1/3 Hz), which is equivalent to the spatial interval, Δx_2 , of 30m. The track spacing, Δx_1 , is 30m and the flying (constant) altitude, h , is at 10m above maximum ground elevation. All these parameters are summarized in Table 5.1.

According to the work done by Vasco and Taylor (1991), the gravitational gradients are simulated by assuming that the GGSS data contain the long and medium wavelength information corresponding to subsurface structures below the geoid (the

¹ NED is designed to provide National elevation data in a seamless form with a consistent datum, elevation unit, and projection (one can visit the USGS's website at <http://www.usgs.gov/> for more information and data availability).

The statistical description of terrain				
mean (m)	max.(m)	min.(m)	std.(m)	rms. (m)
302.2	346.2	263.3	13.2	302.7
NOTE:				
<ul style="list-style-type: none"> • The area size of about 36x36 km² covering the latitudes of 33.7-34.1°N and the longitudes of 261.3-261.6°E • The average density of 2.670 g/cm³ • Velocity of helicopter, $v = 10\text{m/s}$ • Altitude of helicopter = 355m altitude (10m above the maximum ground) • Sampling interval, $\tau = 3$ seconds (sampling frequency = $\Delta_f = 1/\tau$ Hz) • Track spacing, $\Delta x_1 = 30\text{m}$ 				

Table 5.1: The survey area of airborne gradiometry and a prospective survey plan.

nominal gravitational gradients are excluded: $\Gamma_{33} \approx 3080\text{E}$, $\Gamma_{22} \approx 1540\text{E}$, and $\Gamma_{11} \approx 1540\text{E}$). The simulated observation data of six gradients are obtained as follows. Since the GGSS data set is adversely affected by noisy platform accelerations and gradiometer outputs, and by poor navigation, we interpolate all data tracks onto a 30mx30m grid. Then, we perform downward continuation from the GGSS altitude of 1000m to 355m by, first, taking the Fourier transform of the data and then multiplying all data by the attenuation factor $\exp\{-2\pi \frac{h}{|\lambda|}\}$, where h = altitude and λ = wavelength.

Because the high frequency noises in the GGSS data set are amplified in downward continuation, only the gradient signals with longer than 5km wavelengths are used. The shorter wavelength signals are removed before downward continuation by applying the low-pass filter², L , with the cut-off frequency, f_c , of 1/5000 [cyc/m] to the GGSS data in the frequency domain.

² The low-pass filter with the cut-off frequency, f_c , of 1/5000 [cyc/m]:

$$L(f) = \begin{cases} 1 & f_1 \text{ and } f_2 \leq f_c \\ 0 & \text{otherwise} \end{cases}$$

$$f_1 = \frac{i}{N_1 \Delta x_1}; f_2 = \frac{j}{N_2 \Delta x_2}; -\frac{N_1}{2} \leq i \leq \frac{N_1}{2} - 1; -\frac{N_2}{2} \leq j \leq \frac{N_2}{2} - 1 \quad (\text{a})$$

For the remaining part of the data simulation, the gradient signals having wavelengths shorter than 5km, which are assumed to reflect the topographic masses, are generated from finite elements of flat-top rectangular prisms using the 30m DEM³ (the number of prisms for an integration area is limited to 20 on either side of the central computation point; and the density is a constant of 2.670 g/cm³). They are added to the corresponding GGSS data in the frequency domain⁴. Finally, we include 1E/√Hz zero-mean white Gaussian noise along the track or 300 E²/(cyc/m)² at a grid point⁵.

Figure 5.1 shows an example of data simulation in the survey area using the 1987 GGSS data set at an altitude of 1000m above the geoid, then downward continued to 355m and supplemented by the gravitational gradients computed at 355m altitude using 30m rectangular topographic prisms. Also added is 1 E/√Hz white noise (see

where N_1 and N_2 are the number of tracks and the number of points along track, and Δx_1 and Δx_2 are the track spacing and the spatial sampling interval, respectively; $N_1 = 1200$ tracks, $N_2 = 1200$ data points along track, $\Delta x_1 = 30\text{m}$ and $\Delta x_2 = 30\text{m}$.

³ The gradient signal using the 30m DEM, where the number of prisms for an integration area is limited to 20 on either side of the central computation point (m, n) ,

$$\Gamma_{m,n}^{DEM} = \Gamma^{DEM}(m\Delta x_1, n\Delta x_2) = \sum_{p=-10}^{+10} \sum_{q=-10}^{+10} \Gamma_{m+p, n+q}^{prism} \quad \text{for } m = 0 \dots N_1 - 1 \text{ and } n = 0 \dots N_2 - 1 \quad (\text{b})$$

where the 30mx30m rectangular prism Γ^{prism} is computed at 355 m altitude using the equation of (2.15) and (2.16) in Chapter 2.

⁴ The data simulation, $\Gamma_{ij}(f_1, f_2)$, at the spatial frequency (f_1, f_2) with $(i, j) = (1, 2, 3)$ is

$$\Gamma_{ij}(f_1, f_2) = \Gamma_{ij}^{GGSS}(f_1, f_2) + H(f_1, f_2)\Gamma_{ij}^{DEM}(f_1, f_2) \quad (\text{c})$$

where

$\Gamma_{ij}^{GGSS}(f_1, f_2)$ = the Fourier transform of GGSS data of component (i, j) multiplied by the downward continuation factor $\exp\{-2\pi \frac{h}{|\lambda|}\}$ with $h = 1000-355 = 645\text{m}$ and $\lambda = 1/\sqrt{f_1^2 + f_2^2}$.

$\Gamma_{ij}^{DEM}(f_1, f_2)$ = the Fourier transform of the (i, j) gradient component using the 30m DEM

$H(f_1, f_2) = 1 - L(f_1, f_2)$ (i.e., the high-pass frequency with the cut-off frequency f_c)

⁵ We assume the helicopter is moving with the constant velocity, v , of 10m/s. Thus, the along-track gradient (white noise) variance is computed at a point along the track using the sampling frequency $\Delta_f =$

1/3 Hz is $\sigma_w^2 = 1\text{E}^2/\text{Hz} \times 1/3 \text{ Hz} = 1/3 \text{ E}^2$. Furthermore, we simply convert the along-track 1E/√Hz noise at points along the track to the noise at grid points by multiplying the product of $(1\text{E}^2/\text{Hz}) \times (10\text{m/s})$ with the track spacing $\Delta x_1 = 30\text{m}$; we obtain the 300 E²/(cyc/m)² noise at grid points. In addition, the 1E/√Hz zero mean white Gaussian noise is generated using the MATLAB function

“normrnd(mean, $\sigma_w, m, n)$ with mean = 0E, $\sigma_w = 1/\sqrt{3} \text{ E}$, m (the number of tracks) = 1200 and n (the number of points along track) = 1200”.

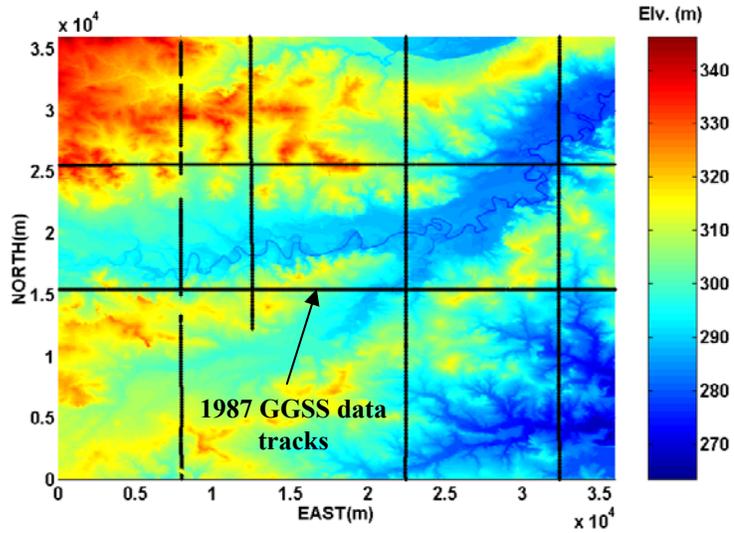
Footnotes); we call this the full geological background plus $1E^2/\text{Hz}$ white noise. The observation, finally, are obtained by adding the gradient signals due to the $2\text{m} \times 2\text{m} \times 1000\text{m}$ target anomaly to the simulated background. Observe that the gradient signals due to the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomalies are not visible in the observations. Table 5.2 lists the root-mean-square (rms) values of the anomaly at 30m below the constant flight altitude, compared with the values of full geological background plus $1E^2/\text{Hz}$ white noise.

Figure 5.2 shows the (azimuth-averaged) empirical power spectral densities (psd's) of the simulated Γ_{33} gradient component at 355m altitude and its corresponding psd models. The psd of the observation (red-dot line) has large amplitudes at the long wavelengths; its trend quickly decreases and then becomes flat at short wavelength (about 300m and shorter) (high frequencies). The magenta-dot line indicates the empirical psd of the residual gradients, $\delta\Gamma_{33}$, after the removal of the gradient field model⁶, $\Gamma_{33}^{\text{model}}$, (or trend surface) (i.e., $\delta\Gamma_{33} = \Gamma_{33}^{\text{Observation}} - \Gamma_{33}^{\text{model}}$). The blue-dot line is the empirical psd of the GGSS gradient plus $1E^2/\text{Hz}$ white noise; as indicated, wavelengths of 5km and shorter are essentially removed but noise at all frequencies still remains.

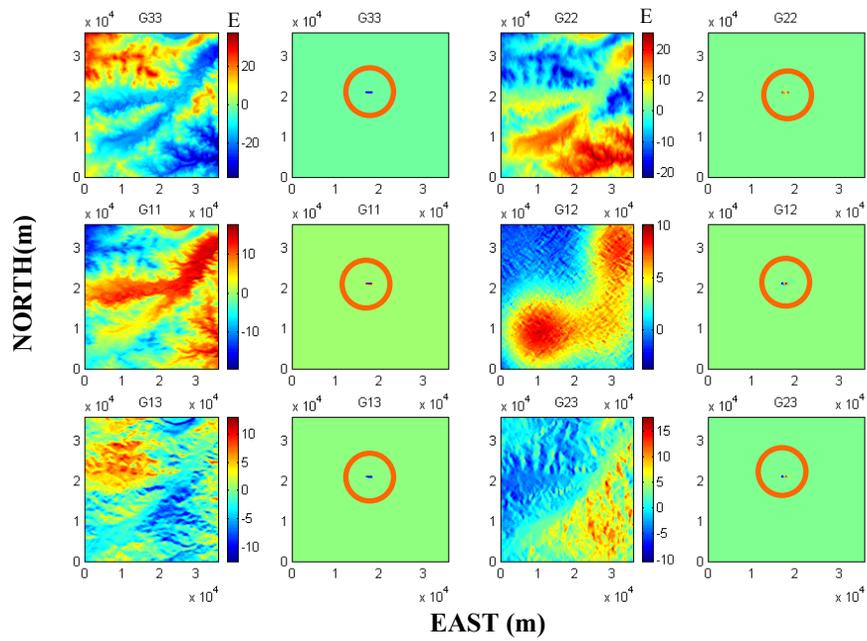
The psd models (smooth curves in Figure 5.2) of all types of gradient observations (except the $1E^2/\text{Hz}$ white noise) at 355m altitude are constructed using the reciprocal distance function models of the disturbing potential (eqs. (B-49) and (B-52) in Jekeli (2003)), based on the 30m DEM over Texas/Oklahoma area. Appendix A lists the corresponding values of the parameters. In this study, we shall assess the limitation of the matched filters that use the psd models, compared with the matched filters that use the empirical psd's⁷ obtained from the simulated observations (that exclude the gradient signals due to the anomaly). One may consider these types of filters (with empirical psd's) as ideal filters. Furthermore, the combination of individual matched filters are assessed. We shall consider no correlation between gradient sensor errors under the white noise assumption. The correlated background field "noise" is accounted for in the filter (unless other specified) using equation (A.1).

⁶ The gradient field model, $\Gamma_{ij}^{\text{model}}$ for $(i, j) = (1, 2, 3)$, is generated from 120m rectangular topographic prisms at the altitude of 355m using the 120m x 120m DEM (the number of prisms over an integration area is limited to 20 on either side of the central computation point, and the density is a constant 2.670 g/cm^3). Then, the gradients at 120m x 120m grids are interpolated onto 30mx30m grids (we use a bicubic spline interpolation function in MATLAB tools for the computation).

⁷ The characteristics of such ideal psd's are similar to the empirical psd's of the observations (wiggly curves) in Figure 5.2.



(a)



(b)

Figure 5.1: (a) The survey area of GGSS airborne gradiometry (b) an example of data simulation by combination of the 1987 GGSS data set and 30m DEM, compared with six gradient components due to the 2m x 2m x 1000m anomaly (see orange circles).

Full geological background plus 1E²/Hz white noise						
Statistics	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
mean	-4.1710	2.4256	1.7451	3.4673	-0.5019	1.7312
med	-5.3099	2.2716	1.4840	3.8059	-0.6960	1.4717
min	-39.0344	-22.3920	-20.4062	-5.0273	-13.8893	-11.5058
max	37.5898	26.5314	18.8898	11.3453	13.9873	18.8403
std	13.2361	9.0212	6.6315	2.9447	3.7420	4.0559
rms	13.8777	9.3416	6.8573	4.5490	3.7755	4.4099
The 2mx2mx1000m anomaly (azimuthal orientation = 90°)						
mean	0.000	0.000	0.000	0.000	0.000	0.000
med	0.000	0.000	0.000	0.000	0.000	0.000
min	-0.9275	-0.1277	-0.1170	-0.1450	-0.4449	-0.3799
max	0.1052	0.1714	0.8881	0.1450	0.4449	0.3799
std	0.0262	0.0028	0.0257	0.0027	0.0188	0.0040
rms	0.0262	0.0028	0.0257	0.0027	0.0188	0.0040
Signal energy	27.7275	0.3274	26.6476	0.2892	14.2964	0.6333
$\text{rms} = \sqrt{\sum_{i=0}^{N-1} s_i^2 / N} ; \text{Signal energy} = \sum_{i=0}^{N-1} s_i^2 \text{ with } N = 1200 \times 1200$ density contrast of anomaly = -1.500 g/cm ³						

Table 5.2: The statistical description of full geological background plus 1E²/Hz white noise and the 2m x 2m x 1000m anomaly at 30m depth from a flight line; the unit of all values is Eötvös.

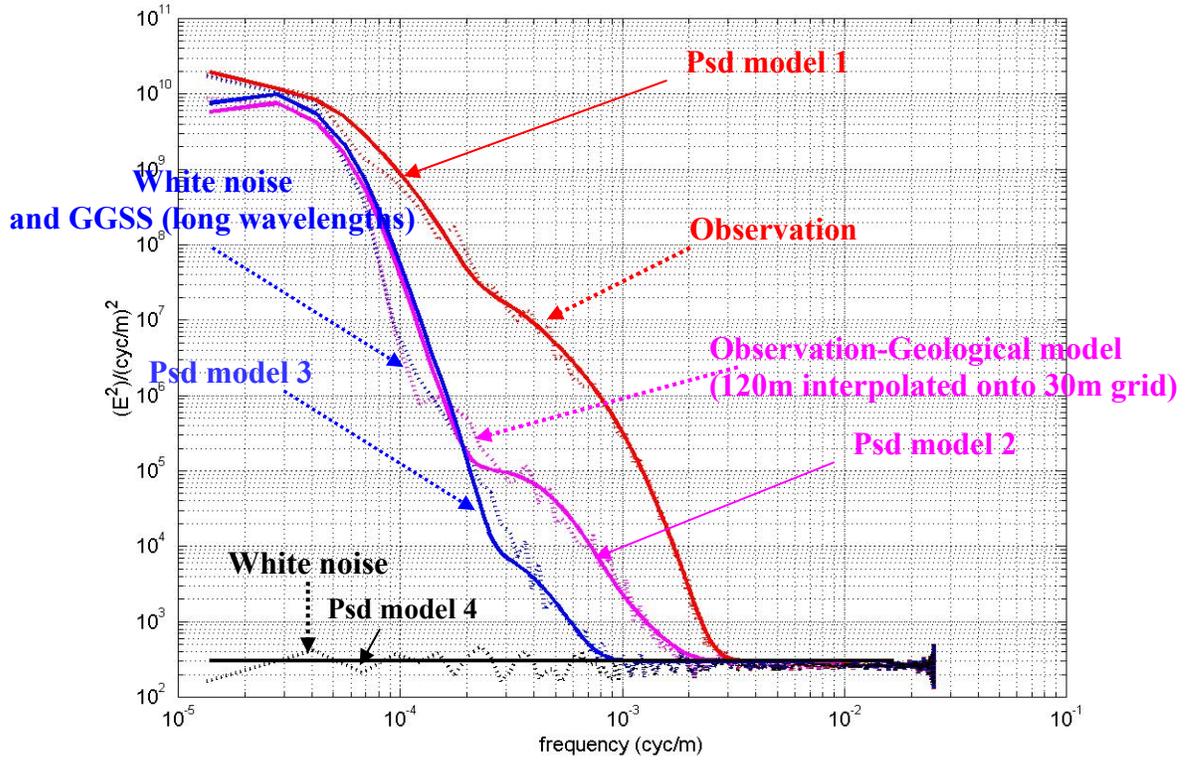


Figure 5.2: The (averaged-azimuth) empirical psd's of the Γ_{33} component and its corresponding psd models (for other components, the same parameters of the psd models are used, and the corresponding psd's agree with the empirical psd's of those components (not shown here)).

5.2 CASE I: White noise environment

In an example of detectability using the matched filter for the case of $1E/\sqrt{\text{Hz}}$ white noise ($\sigma_w^2 = 1/3 E^2$ for $v = 10\text{m/s}$ and $\Delta_f = 1/3 \text{ Hz}$), we consider the gravitational gradients due to the target anomaly of 2m-width x 2m-height x 1000m-long, at the depth (D) of 30m below a flight line, shown in Figure 5.3a.

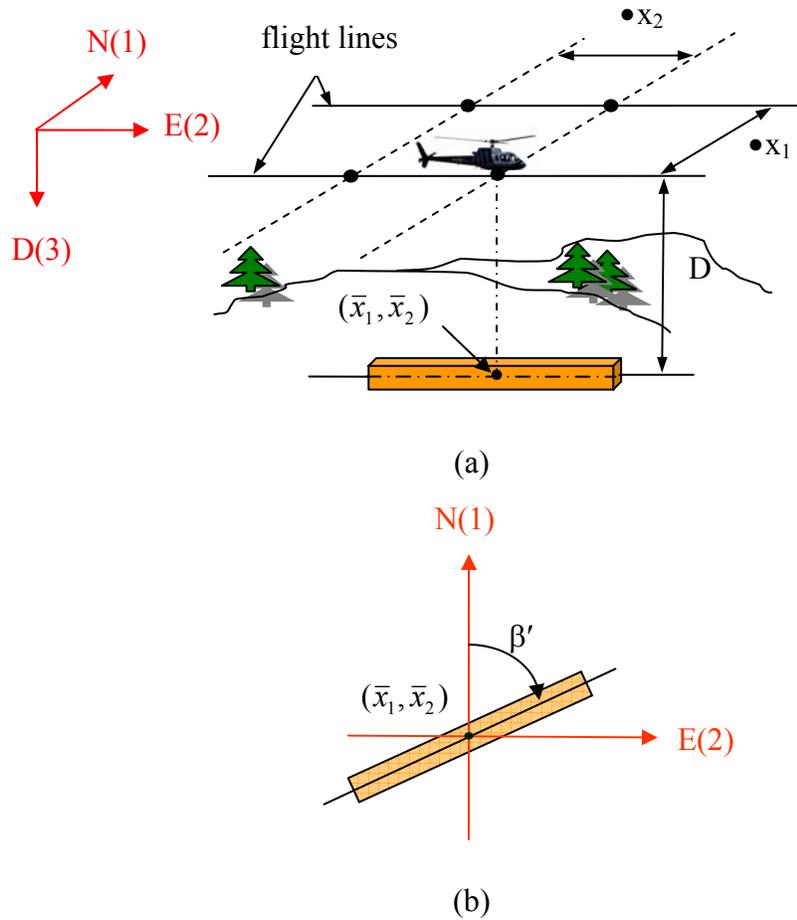


Figure 5.3: (a) The gradiometry survey; (b) The 10m x 5m x 1000m anomaly at the azimuthal orientation of β' .

The center of anomaly is at $\bar{x}_1 = 18000.0\text{m}$ and $\bar{x}_2 = 18000.0\text{m}$ of the survey grid. The anomaly is rotated by 90 degrees about the vertical axis (i.e., $\beta' = 90^\circ$), positive in clockwise direction, with respect to the N-direction, shown in Figure 5.3b. Figure 5.4 shows the observations of all six gradient components in $1E^2/\text{Hz}$ white noise. It should be noted that the gradient signals due to the anomaly are not detectable visually.

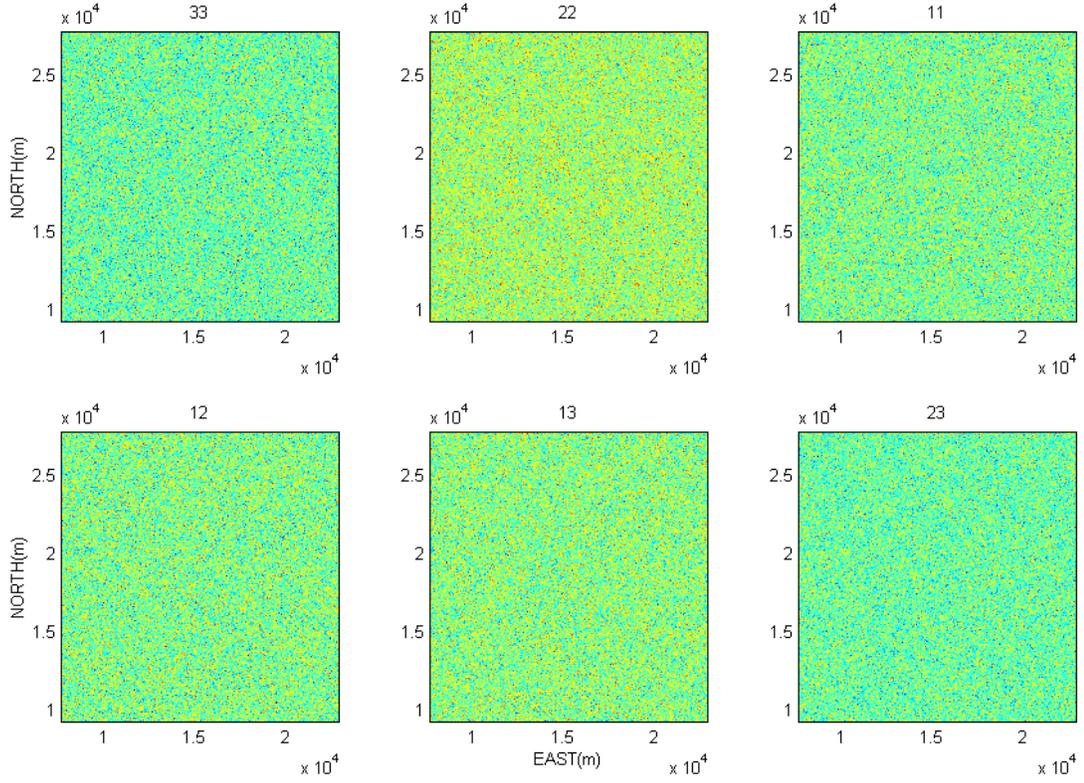


Figure 5.4: Six gradient components due to the 2m x 2m x 1000m anomaly in $1E^2/Hz$ white noise.

Suppose we know the parameters of the target anomaly except its location, i.e., we know its size, depth, and orientation. Based on (3.12), we design the two-dimensional matched filter function h_{ij} for the (i, j) gradient component, with $(i, j) = (1, 2, 3)$, as a reversed replica of the actual signal such that

$$h_{ij}(x_1, x_2) = \frac{1}{\sigma_w^2} s_{ij}(-x_1, -x_2) \quad (5.1)$$

where (x_1, x_2) is an arbitrary location. The unit of h is defined as $1/E$ so that the filter output is unitless for convenience. The transfer function of h_{ij} is:

$$H_{ij}(f_1, f_2) = \frac{1}{\sigma_w^2} \tilde{S}_{ij}^*(f_1, f_2) \quad (5.2)$$

the unit of which is $\left[\frac{1}{E} \right] \left[\frac{m}{\text{cyc}} \right]^2$. Figure 5.5 shows the pairs of the matched filter functions and their transfer functions for individual gradient components.

If the filter matches the gradient signal, it yields the maximum signal-to-noise ratio

$$\text{SNR}_{\max} = \frac{E_s}{\Delta x_1 \Delta x_2 \sigma_w^2} \quad (5.3)$$

where, using Parseval's theorem as (3.10) in Chapter 3, the signal energy E_s is

$$E_s = \int_{-f_{N_2}}^{+f_{N_2}} \int_{-f_{N_1}}^{+f_{N_1}} \left| \tilde{S}_{ij}(f_1, f_2) \right|^2 df_1 df_2 = \Delta x_1 \Delta x_2 \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} s_{ij}(x_{1,m}, x_{2,n})^2 \quad (5.4)$$

The symbol “ \tilde{S}_{ij} ” is the (periodic) Fourier transform of s_{ij} and Nyquist frequencies f_{N_1} and f_{N_2} are $1/2\Delta x_1$ and $1/2\Delta x_2$ respectively. It should be noted in (5.3) with (5.4) that a matched filter performs best for detecting a signal in the presence of noise only if the gradient waveform is known completely. The location of the maximum (highest peak) output of the matched filter corresponds to the maximum SNR.

As regards (3.14) but applied to two dimensions, the matched filter yields the highest peak $y_h(\bar{x}_1, \bar{x}_2)$ at the point (\bar{x}_1, \bar{x}_2) which is the center location of the anomaly to be detected such that

$$y_h(\bar{x}_1, \bar{x}_2) = \frac{1}{\sigma_w^2} \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} s_{ij}(x_{1,m} - \bar{x}_1, x_{2,n} - \bar{x}_2)^2 + \frac{1}{\sigma_w^2} \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} s_{ij}(x_{1,m} - \bar{x}_1, x_{2,n} - \bar{x}_2) n(x_{1,m}, x_{2,n}) \quad (5.5)$$

Clearly, the first term on the right-hand side of (5.5) equals SNR_{\max} . Practically, it should be noted that the highest peak may be larger or smaller than SNR_{\max} since the second term of (5.5) can be either positive or negative.

The noise in the second term of (5.5) may cause the highest output peak to be not at the true location, i.e., the point (\bar{x}_1, \bar{x}_2) , if the gradient signal is embedded in a strong noise background (see the examples in Table 5.2 where the filters of the components Γ_{22} , Γ_{12} , and Γ_{23} have small rms values compared with those of full geological background plus $1E^2/\text{Hz}$. The consequence of small signals relative to strong noise background will be discussed later).

In Figure 5.6, the matched filters for gradient components maximize the SNR at the true location of the anomaly, see Table 5.3. The locations of the output peaks

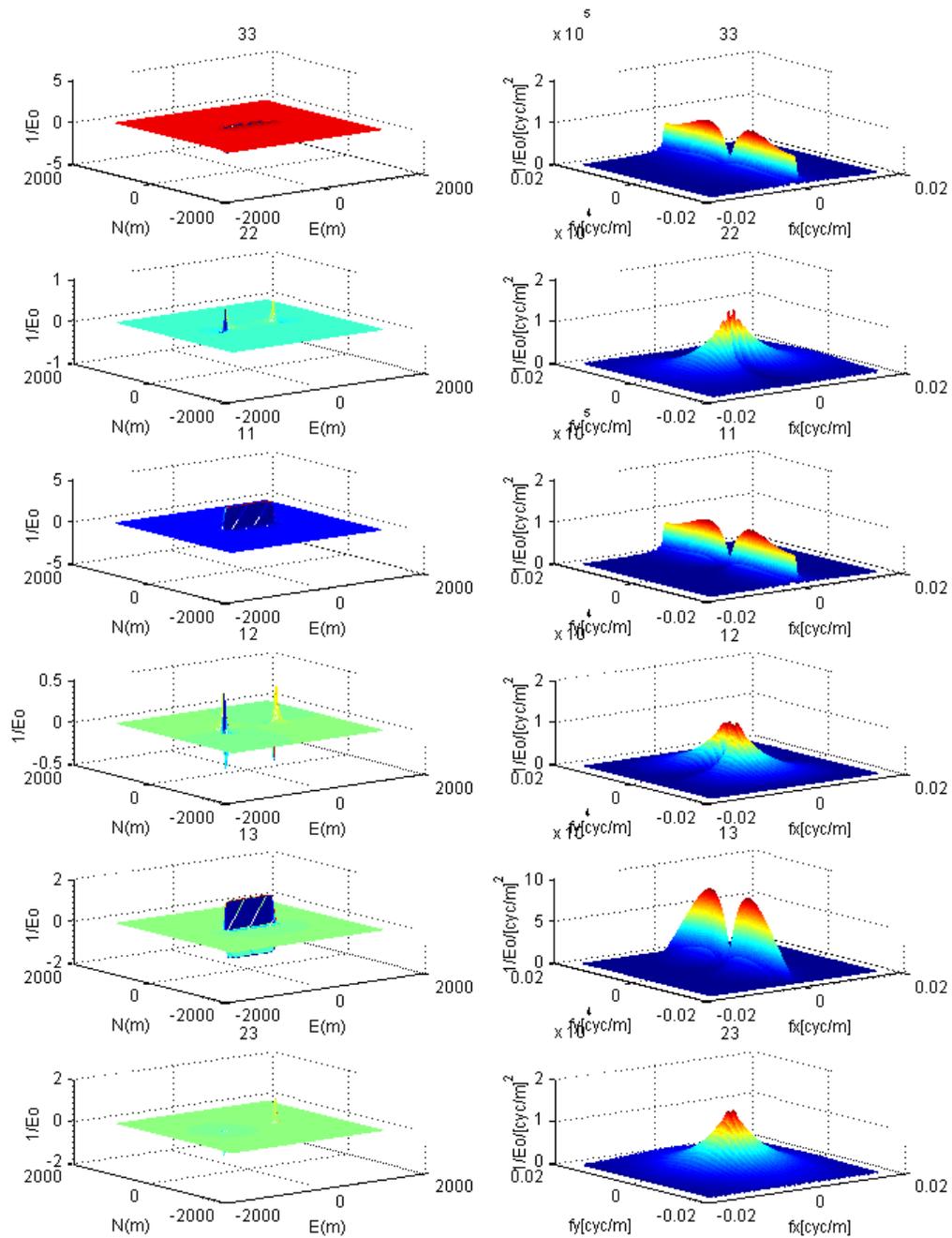


Figure 5.5: The matched filter functions and the transfer functions (using Psd model 4) for the individual gradients due to the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90° and 30m depth in 1E²/Hz white noise.

	Gradient component					
	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
ΔN (m)	0.0	5400.0	0.0	4830.0	0.0	-6030.0
ΔE (m)	-90.0	-6030.0	30.0	-7230.0	0.0	4440.0

Table 5.3: Signal detection in $1E^2/\text{Hz}$ white noise environment; ΔN and ΔE are the difference of the location of the highest peak minus the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly.

(symbol “ \times ”) for the Γ_{33} and Γ_{11} components are shifted from the true location (symbol “+”) by 90m westward and 30m eastward, respectively. These errors are due to the gradient signals corrupted by the white noise. This noise has a much greater effect on the filters for components Γ_{22} , Γ_{12} , and Γ_{23} , which cannot locate the anomaly at all.

According to (3.29), the matched filter output is indeed the test statistic (or sufficient statistic) (which contributes towards a means of making a decision about the acceptance of H_0 (i.e., if H_0 is true, the gradient signal due to the anomaly is present)). The performance of signal detection can be statistically described by defining an a priori POM such that it is a useful confidence level related to the decision to accept H_0 in (3.1). The corresponding threshold can be computed from (3.43) for a given POM. If the highest peak after matched filtering exceeds the threshold, then the signal is presumed to be presented.

Table 5.4 shows the detectability performance of the matched filters for given $\text{POM} = 0.001, 0.050, \text{ and } 0.010$. The Γ_{33} , Γ_{11} , and Γ_{13} matched filters yield the highest output peaks at or close to the true location. These outputs imply high performance of the filters. For instance, $\text{POF} = 0.000$ for the given POM 's, for these components. H_0 although, in fact, the highest peak output occurs at a point different from the true location (see also Figure 5.6) by 4830m northward and 7230m westward (taken from Table 5.3). In this study, although the threshold seems to be useless because the highest peak is selected regardless of it, it is a preliminary step that statistically describes the possible candidate(s).

The POF obtained from (3.44) is a useful probability when making a decision to accept H_0 . It should be noted, however, that POF increases if we decrease POM . Thus, the POM should be carefully chosen so that the corresponding POF is useful at most to aid the decision. In Table 5.4, the POF 's of the component Γ_{12} indicate that a chance of

no signal present is high; POF = 0.983, 0.911, and 0.762 for the given POM's 0.001, 0.010, and 0.050, respectively.

For the case of $1E^2/Hz$ white noise environment, the shape of the gradient signal does not affect the performance of the matched filter if Psd model 4, which is flat and constant (shown in Figure 5.2), is used. Therefore, the gradient signals render the same detection performance if their energies are similar, see (5.3) with (5.4). The next example will show how correlated noises affect signal detection by the matched filters to which Psd model 1 and empirical psd's are applied.

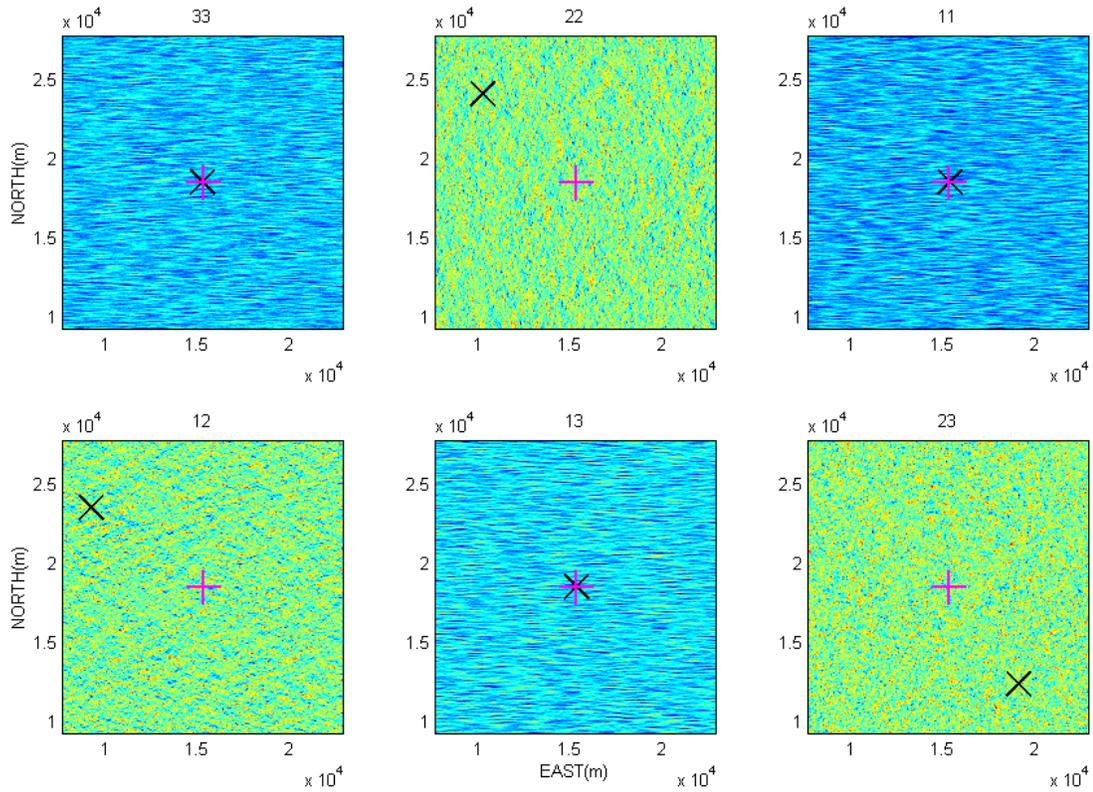


Figure 5.6: The matched filter outputs (for $1E^2/Hz$ white noise): “+” = the true location of the 2m x 2m x 1000m anomaly to be detected; “x” = the location where the highest peak occurs.

Grad.	d^2	y_{\max}	$\alpha = 0.001$		$\alpha = 0.010$		$\alpha = 0.050$	
			η'	β	η'	β	η'	β
Γ_{33}	76.231	84.516	49.250	0.000	55.920	0.000	61.870	0.000
Γ_{22}	0.963	4.577	-2.070	0.983	-1.320	0.911	-0.651	0.746
Γ_{11}	73.008	81.431	46.603	0.000	53.130	0.000	58.953	0.000
Γ_{12}	0.867	4.331	-2.011	0.985	-1.299	0.918	-0.665	0.762
Γ_{13}	39.197	50.090	19.850	0.001	24.632	0.000	28.899	0.000
Γ_{23}	1.881	6.210	-2.357	0.957	-1.310	0.830	-0.375	0.608

Table 5.4: Performance of matched filter in white noise environment; $d^2 = \text{SNR}_{\max}$, y_{\max} = the highest peak output, β = POF, η' = the computed threshold with a given POM = α and d^2 (unitless).

5.3 CASE II: Correlated noise environment

In practical gradiometry, geological background and instrumental noises are the primary noise sources that corrupt the gradient signal due to an anomalous source. As described in Section 5.1, Figure 5.2 shows the power spectrum obtained by the simulated gradient field. The red-dot line presents the empirical psd of the observation, which does not flat. We use Psd model 1 (red-solid line) to show the result of detection by the matched filtering technique. As stated in Section 5.1, we shall assess the limitation of the matched filters with the use of the psd models (see also the smooth curves in Figure 5.2), compared to the ideal filters for which the empirical psd's are obtained from the simulated gradient field (without the gradient signal due to the anomaly).

This example shows the anomaly detection by matched filters in a correlated noise environment. The corresponding filters can be expressed and numerically computed using the equations in Section 3.5 in Chapter 3 but extended to the two-dimensional case:

$$h_{ij}(x_1, x_2) = \Delta x_1 \Delta x_2 \int_{-f_{N_2}}^{+f_{N_2}} \int_{-f_{N_1}}^{+f_{N_1}} \frac{\tilde{S}_{ij}^*(f_1, f_2)}{\Phi_n^{a \text{ or } b}(f_1, f_2)} e^{i2\pi(f_1 x_1 + f_2 x_2)} df_1 df_2 \quad (5.7)$$

with a = psd model and b = empirical psd, and its transfer function is

$$H_{ij}(f_1, f_2) = \Delta x_1 \Delta x_2 \frac{\tilde{S}_{ij}^*(f_1, f_2)}{\Phi_n^{a \text{ or } b}(f_1, f_2)} \quad (5.8)$$

which has units of $\left[\frac{1}{\text{E}} \right] \left[\frac{\text{m}}{\text{cyc}} \right]^2$. The SNR_{\max} is

$$\text{SNR}_{\max} = \int_{-f_{N_2}}^{+f_{N_2}} \int_{-f_{N_1}}^{+f_{N_1}} \frac{|\tilde{S}_{ij}(f_1, f_2)|^2}{\Phi_n^{a \text{ or } b}(f_1, f_2)} df_1 df_2 \quad (5.9)$$

What follow are the results of detecting the 2m x 2m x 1000m anomaly by the matched filters using Psd model 1 versus the empirical psd for full geological background plus $1\text{E}^2/\text{Hz}$ white noise. The cases of the anomaly at the azimuthal orientation, β' , of 90° , 45° , and 0° are included. Although we perform matched filtering through the entire area of data simulation (i.e., 36km x 36km), only the filtering output covering the 20km x 20km innermost area of the simulation is used to avoid edge effects and also cyclic convolution errors⁸.

Figures 5.7 and 5.9 show the pairs of the matched filters and their transfer functions using Psd model 1 and the corresponding empirical psd, respectively, where $\beta' = 90^\circ$ is

⁸ Since Fortran 90 codes for this study have a limitation of array memory, we are not able to apply zero-padding to the original signal array (having 1200x1200 arrays); efficient algorithms should be considered for future work. However, cyclic convolution errors can be reduced by considering only a subset of the domain of the computation of the convolution.

considered. The matched filter outputs are shown in Figures 5.8 and 5.10 in sequence. Table 5.5 summarizes the filter outputs and their detection performance. Only the matched filters for the Γ_{33} , Γ_{11} , and Γ_{13} components, either using the psd model or the empirical psd, provide the highest peak at or close to the true location. However, there are no significant differences between the performance of the filters using the psd model and the empirical model although, for instance, the corresponding SNR values are $d^2 = 71.503$ and 249.512 for the Γ_{33} component, respectively. That is $\text{POF} = 0.000$ for the given $\text{POM} = 0.001, 0.010, \text{ and } 0.050$ in both cases. The filters using either the psd model or the empirical psd yield large values of POF when the highest peaks do not locate or are close to the true location. Note that the Γ_{22} , Γ_{12} , and Γ_{23} filters using either the smooth psd or the empirical psd can not locate the anomaly for $\beta' = 90^\circ$ and $\beta' = 0^\circ$, see also Tables 5.7. This is because the gradient signals are relatively small, compared with the noise background; the capability of the filters has its limitation in a strong noise background.

For $\beta' = 45^\circ$ and the component Γ_{11} , the signal is not detectable by the Γ_{11} matched filter using Psd model 1 whereas the filter using the empirical psd is; see Table 5.6. The capabilities of the filters between using Psd model 1 and the empirical psd are significantly different. The latter shows larger magnitudes of the transfer functions (eq. (5.8)) of the matched filters than those using the psd model. For instance, shown in Figure 5.11 versus 5.13, the maximum magnitude of the Γ_{11} transfer function using the psd model is about $5 \times 10^4 [1/E]/[\text{cyc}/\text{m}]^2$ which is smaller than the maximum magnitude using the empirical psd ($= 5 \times 10^6 [1/E]/[\text{cyc}/\text{m}]^2$). This is because noise (empirical) spectrum can contain many larger “spikes”. The matched filter using the smooth psd can not sufficiently represent all these spikes while the filter using the empirical psd can.

By comparing the detection results, shown in Figure 5.12 versus 5.14, the matched filters using the empirical psd outperform the filters using the psd model. Detecting the anomaly by the former is done successfully. In Table 5.6, although $\text{POF} = 0.171$ for the given $\text{POM} = 0.001$, the Γ_{11} matched filter using the psd model mislocates the target anomaly. By using the empirical psd's, the filters provide improved noise suppressing capabilities and can better locate the true location of the anomaly. It is clear that the more accurate the psd modeling, the better is the performance of matched filter.

Figures 5.15 – 5.18 show the case $\beta' = 0^\circ$. Table 5.7 summarizes signal detectability and performance of matched filters. Although the performance of the Γ_{33} matched filter using Psd model 1 is relatively high, compared to other filters, the filter can not locate the target, which can be detected by the filter using the empirical psd. Therefore, this result confirms that a more accurate psd model than Psd model 1 is needed to improve the detectability. In this section, we have not discussed how orientations will affect the detectability of matched filters. More details will be discussed in Section 5.6.

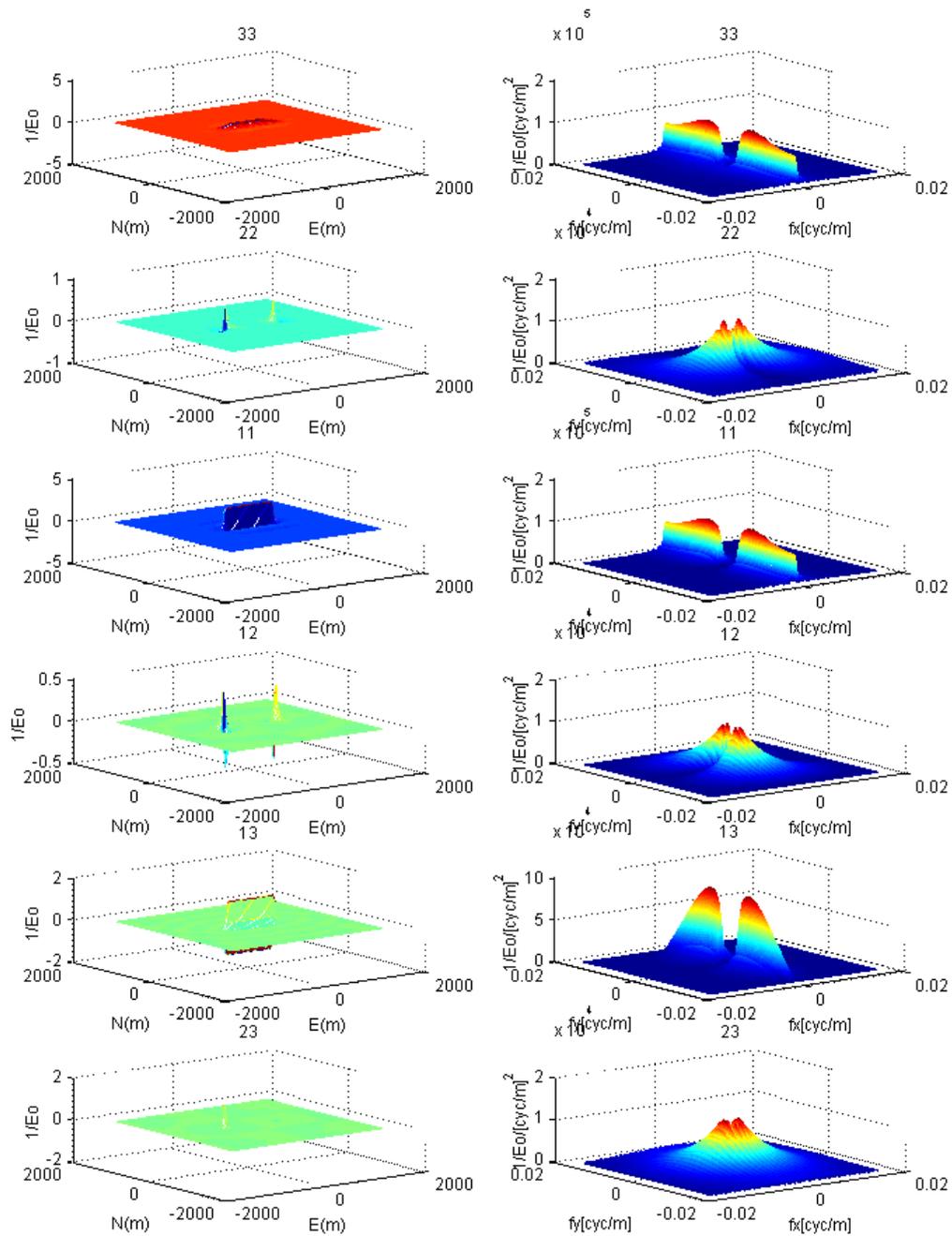


Figure 5.7: The matched filter functions and the transfer functions (using Psd model 1) for the individual gradients due to the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90° and 30m depth in the full geological background plus 1E²/Hz white noise.

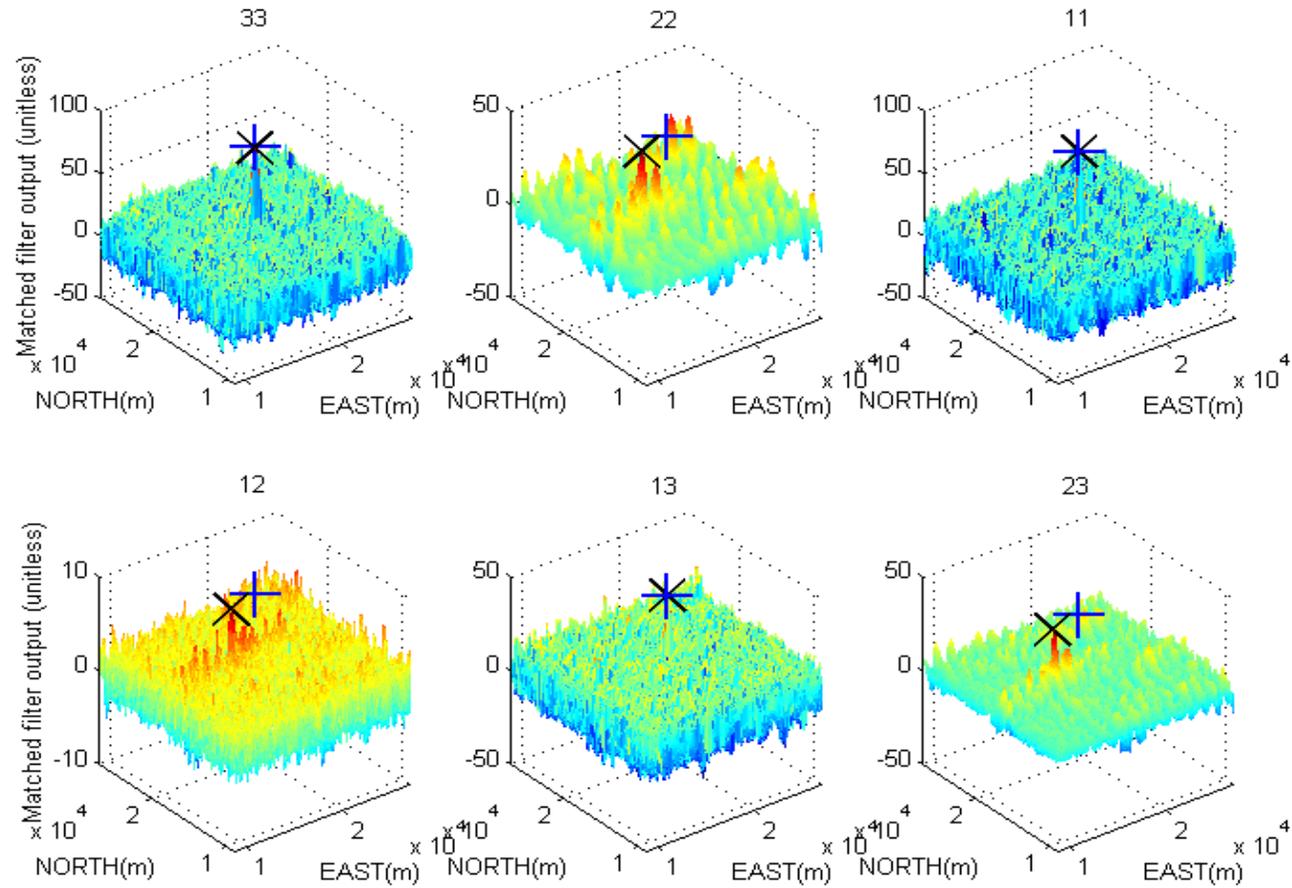


Figure 5.8: The matched filter outputs (using Psd model 1 for the full geological background plus $1E^2/\text{Hz}$ white noise); “x” indicates the location of highest peak; “+” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 90° and 30m depth.

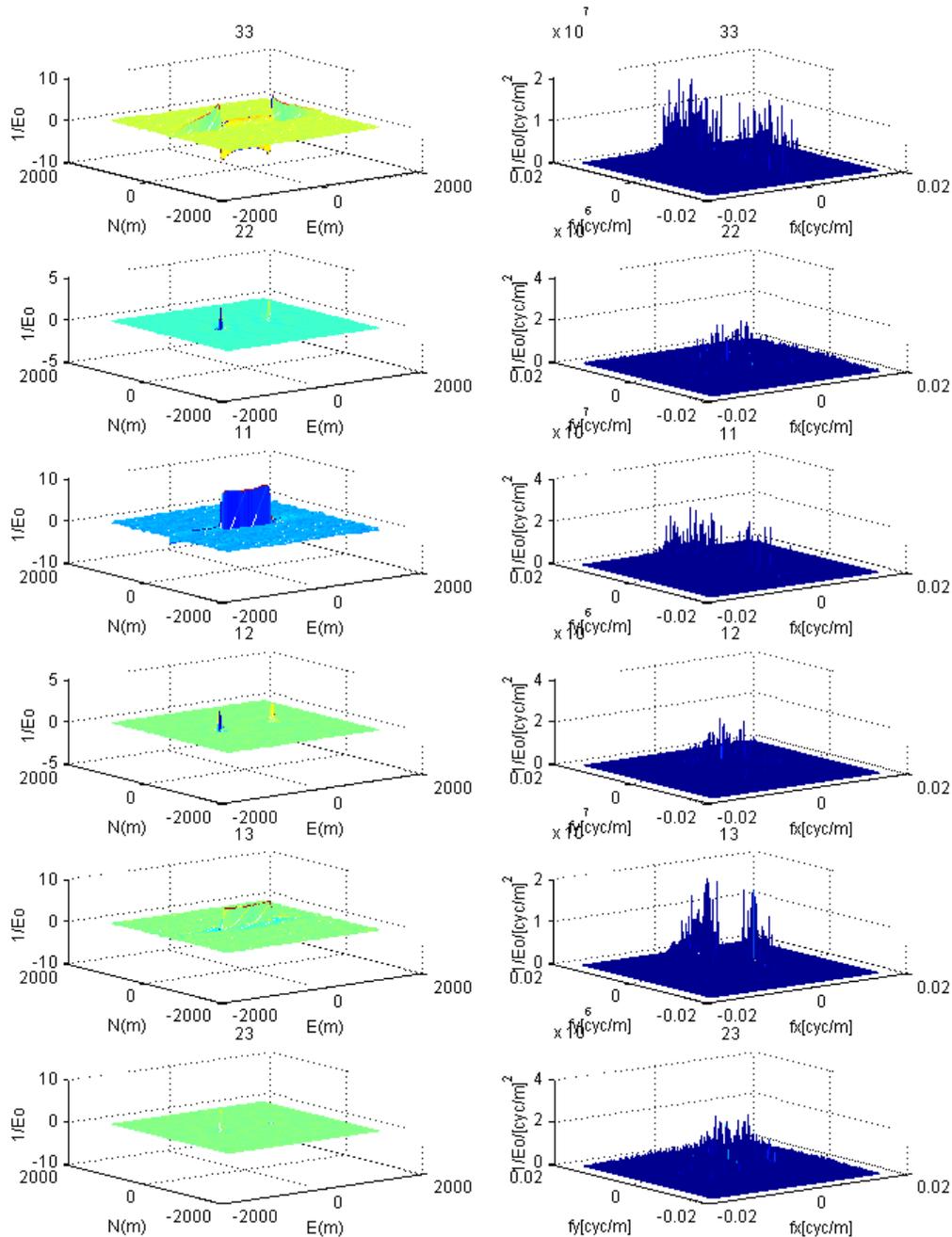


Figure 5.9: The matched filter functions and the transfer functions (using the empirical psd) for the individual gradients due to the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90° and 30m depth in the full geological background plus $1E^2/\text{Hz}$ white noise.

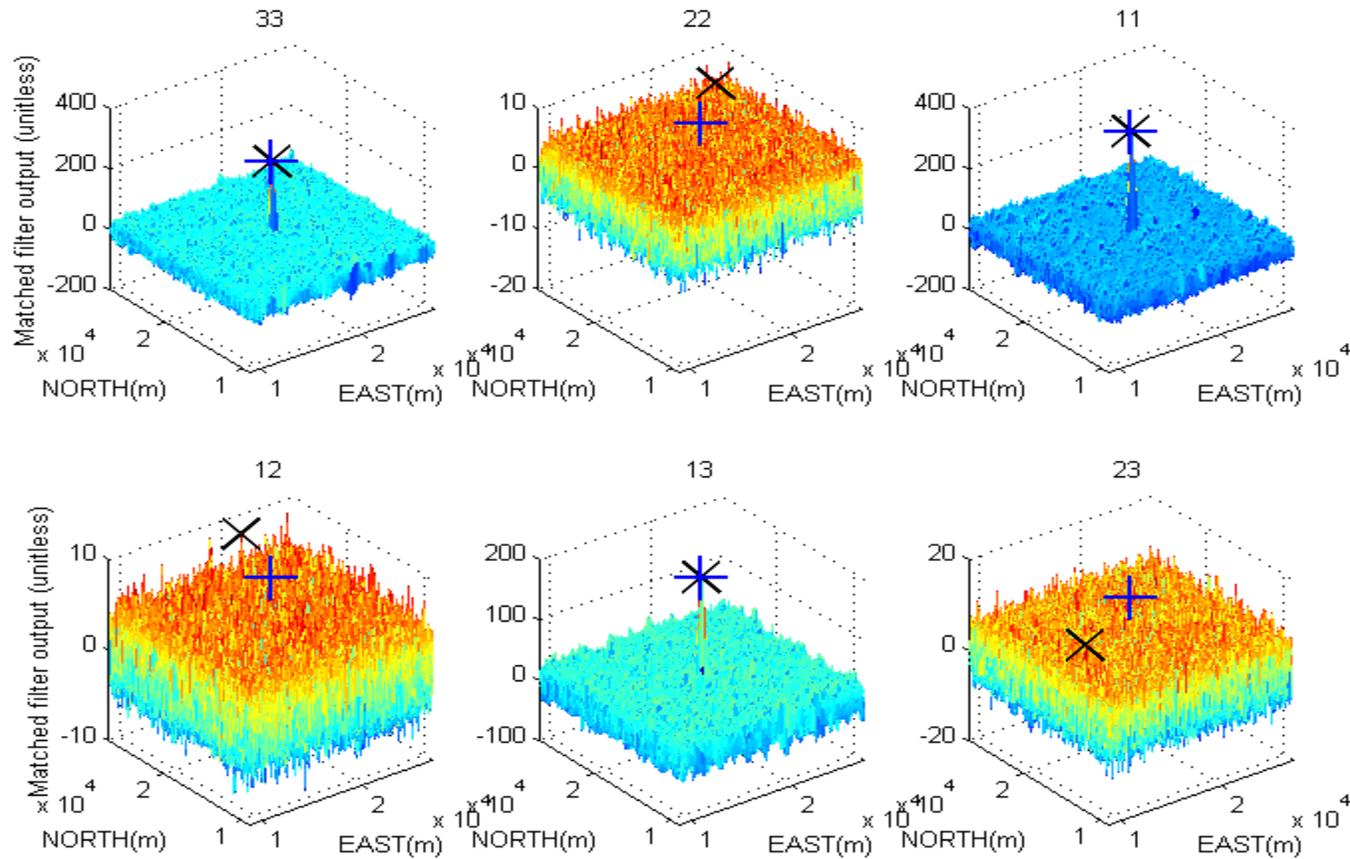


Figure 5.10: The matched filter outputs (using the empirical psd for the full geological background plus $1E^2/\text{Hz}$; white noise) “x” indicates the location of highest peak; “+” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly; at the azimuthal orientation of 90° and 30m depth.

Type	Grad.	ΔN	ΔE	d^2	y_{\max}	POM = 0.001		POM = 0.010		POM = 0.050	
						η'	POF	η'	POF	η'	POF
a	Γ_{33}	0.0	-90.0	71.503	68.124	45.372	0.000	51.832	0.000	57.595	0.000
	Γ_{22}	-720.0	-3150.0	0.941	40.628	-2.057	0.983	-1.316	0.912	-0.655	0.750
	Γ_{11}	0.0	30.0	68.884	67.895	43.236	0.000	49.576	0.000	55.232	0.000
	Γ_{12}	-810.0	-3090.0	0.807	8.049	-1.969	0.986	-1.283	0.923	-0.671	0.772
	Γ_{13}	0.0	0.0	35.893	32.385	17.379	0.002	21.956	0.000	26.039	0.000
	Γ_{23}	-690.0	-3180.0	1.766	32.331	-2.341	0.961	-1.325	0.841	-0.420	0.624
b	Γ_{33}	0.0	0.0	249.512	254.065	200.699	0.000	212.765	0.000	223.530	0.000
	Γ_{22}	4740.0	5070.0	5.102	9.059	-1.878	0.797	-0.153	0.527	1.387	0.267
	Γ_{11}	0.0	0.0	334.154	353.945	277.665	0.000	291.629	0.000	304.086	0.000
	Γ_{12}	7410.0	2670.0	5.034	9.067	-1.899	0.801	-0.186	0.533	1.343	0.275
	Γ_{13}	0.0	0.0	178.949	185.643	137.611	0.000	147.829	0.000	156.946	0.000
	Γ_{23}	-4530.0	-7950.0	9.824	13.829	0.138	0.482	2.532	0.209	4.668	0.068

Table 5.5: Detectability and performance of the matched filter using (a) Psd model 1 and (b) the empirical psd for the full geological background plus $1E^2/\text{Hz}$ white noise; ΔN and ΔE are the difference of the location of the highest peak minus the true location; the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 90° and 30m depth.

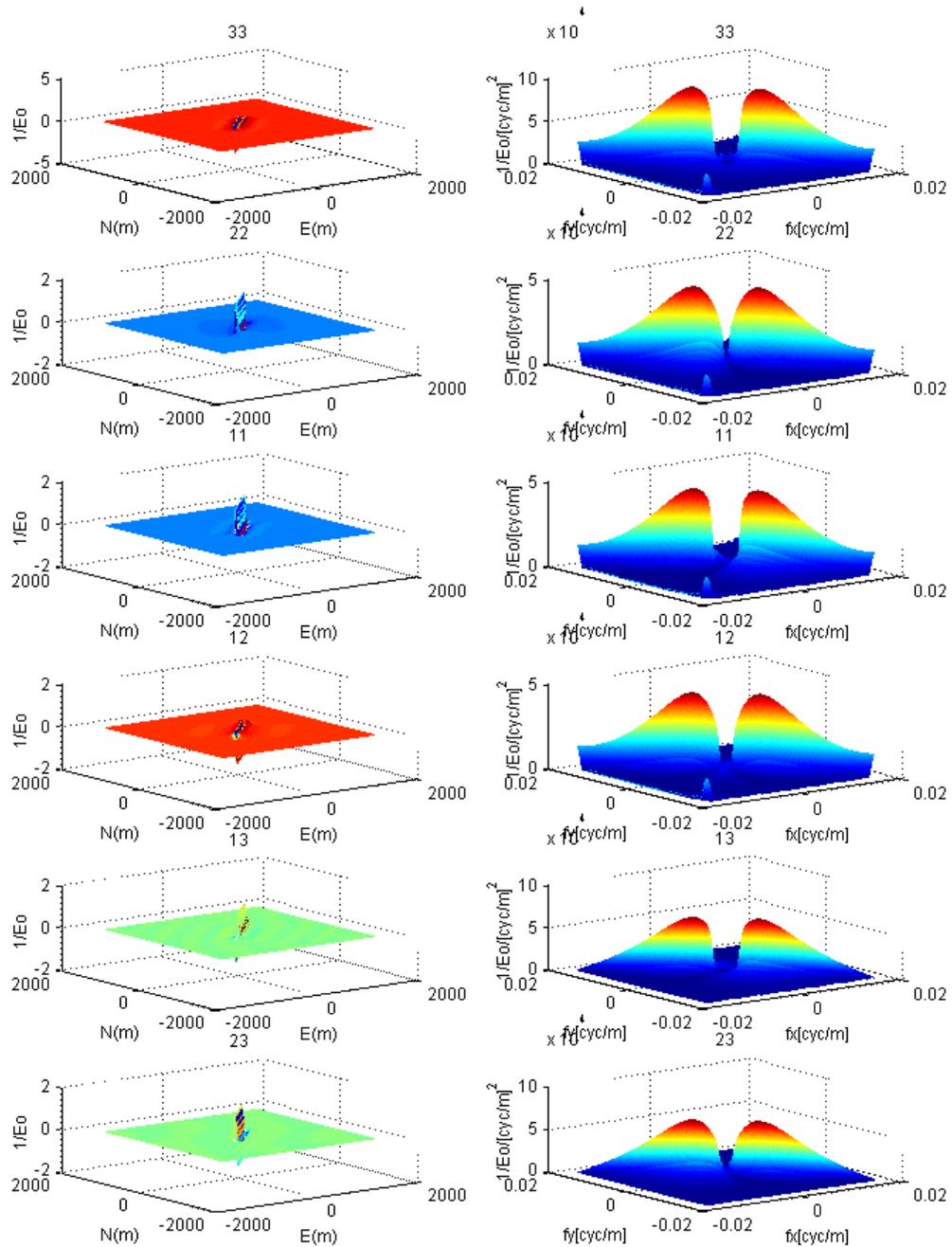


Figure 5.11: The matched filter functions and the transfer functions (using Psd model 1) for the individual gradients due to the 2m x 2m x 1000m anomaly at the azimuthal orientation of 45° and 30m depth in full geological background plus 1E²/Hz white noise.

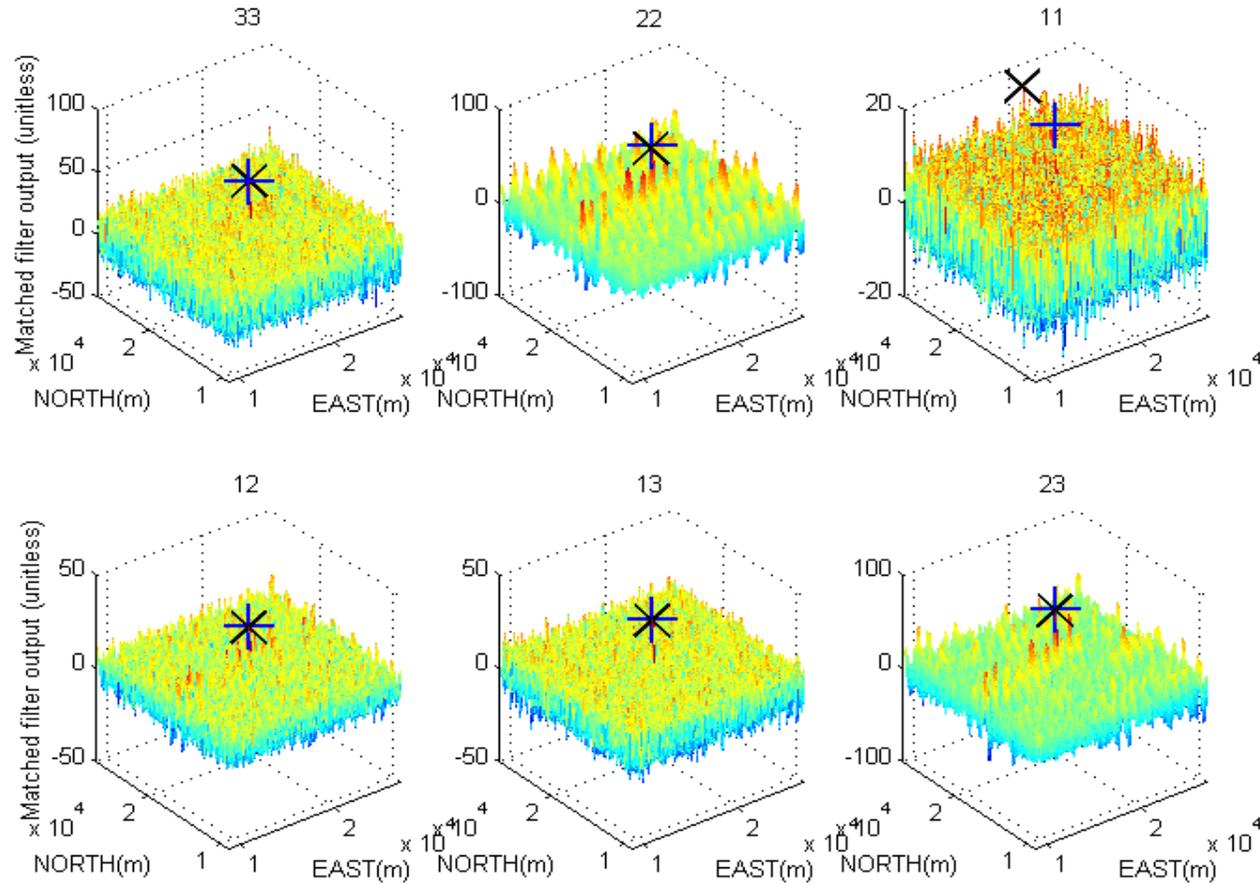


Figure 5.12: The matched filter outputs (using Psd model 1 for the full geological background plus $1E^2/\text{Hz}$ white noise); “x” indicates the location of highest peak; “+” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 45° and 30m depth.

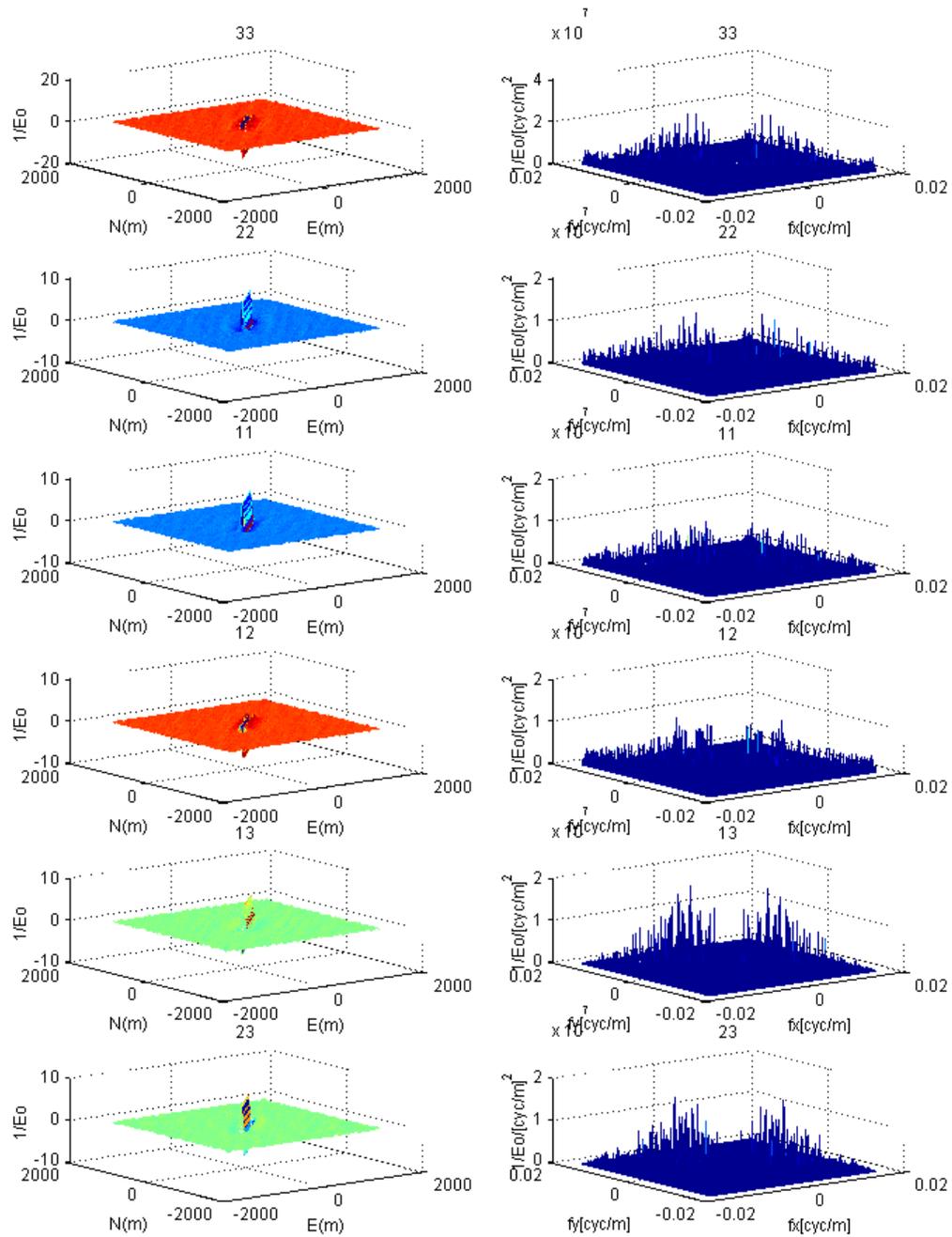


Figure 5.13: The matched filter functions and the transfer functions (using the empirical psd) for the individual gradients due to the 2m x 2m x 1000m anomaly at the azimuthal orientation of 45° and 30m depth in the full geological background plus $1E^2/\text{Hz}$ white noise.

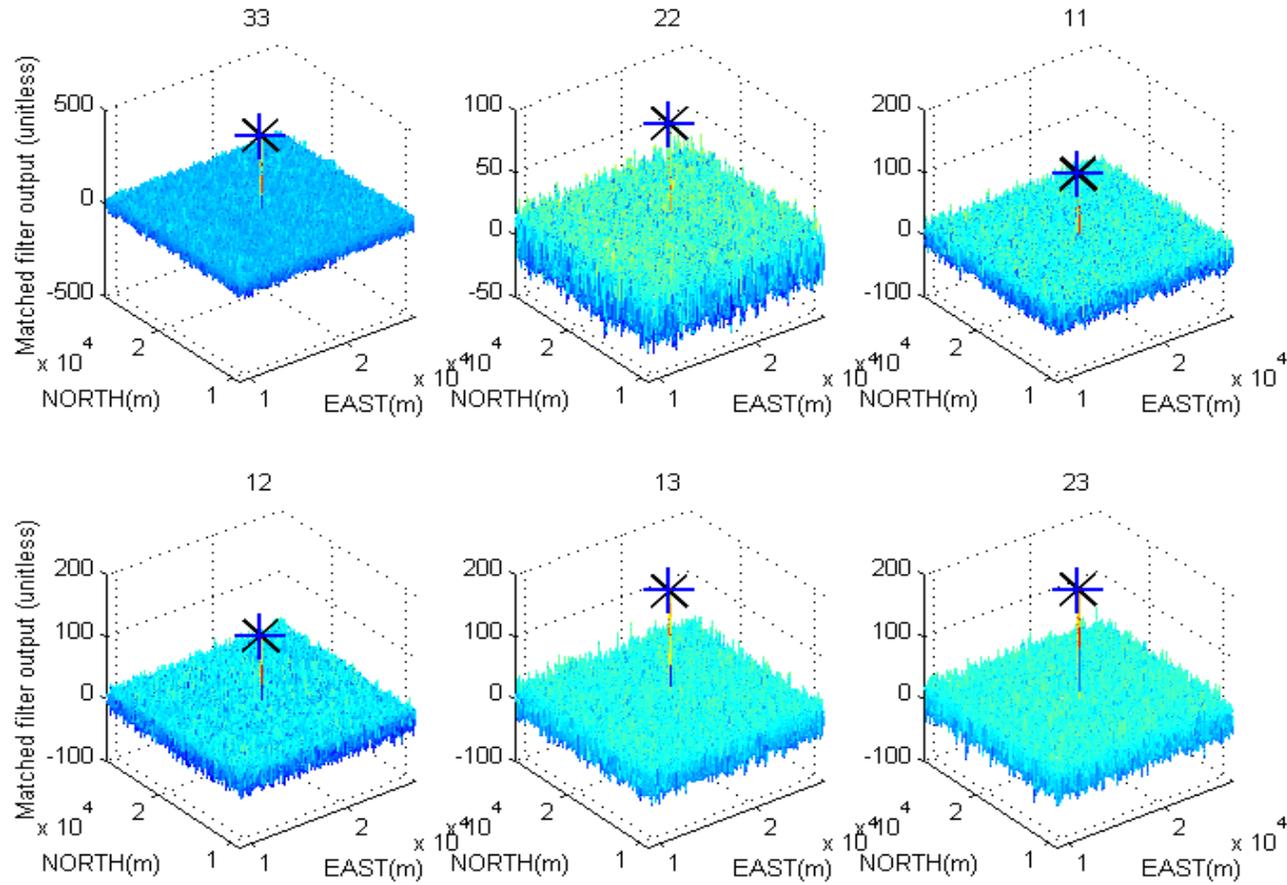


Figure 5.14: The matched filter outputs (using the empirical psd for the full geological background plus $1E^2/\text{Hz}$ white noise); “ \times ” indicates the location of highest peak; “ $+$ ” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 45° and 30m depth.

Type	Grad.	ΔN	ΔE	d^2	y_{\max}	POM = 0.001		POM = 0.010		POM = 0.050	
						η'	POF	η'	POF	η'	POF
a	Γ_{33}	-30.0	-30.0	62.852	50.517	38.353	0.000	44.409	0.000	49.812	0.000
	Γ_{22}	-240.0	-300.0	17.347	71.246	4.476	0.141	7.658	0.033	10.496	0.006
	Γ_{11}	6870.0	1800.0	16.320	18.657	3.836	0.171	6.922	0.043	9.675	0.008
	Γ_{12}	-90.0	-90.0	15.573	27.688	3.378	0.196	6.392	0.053	9.082	0.011
	Γ_{13}	-60.0	-60.0	27.562	31.288	11.339	0.015	15.349	0.002	18.927	0.000
	Γ_{23}	-300.0	-300.0	28.987	74.179	12.349	0.011	16.462	0.001	20.131	0.000
b	Γ_{33}	0.000	0.000	424.216	413.879	360.568	0.000	376.301	0.000	390.338	0.000
	Γ_{22}	0.000	0.000	112.838	96.789	80.012	0.000	88.127	0.000	95.366	0.000
	Γ_{11}	0.000	0.000	114.111	113.064	81.101	0.000	89.261	0.000	96.541	0.000
	Γ_{12}	0.000	0.000	106.288	116.318	74.429	0.000	82.305	0.000	89.331	0.000
	Γ_{13}	-30.000	-30.000	189.514	189.830	146.972	0.000	157.488	0.000	166.870	0.000
	Γ_{23}	-30.000	-30.000	190.639	192.157	147.972	0.000	158.519	0.000	167.929	0.000

Table 5.6: Detectability and performance of the matched filter using (a) Psd model 1 and (b) the empirical psd for the full geological background plus $1E^2/\text{Hz}$ white noise; ΔN and ΔE are the difference of the location of the highest peak minus the true location; the 2m x 2m x 1000m anomaly at the azimuthal orientation of 45° and 30m depth

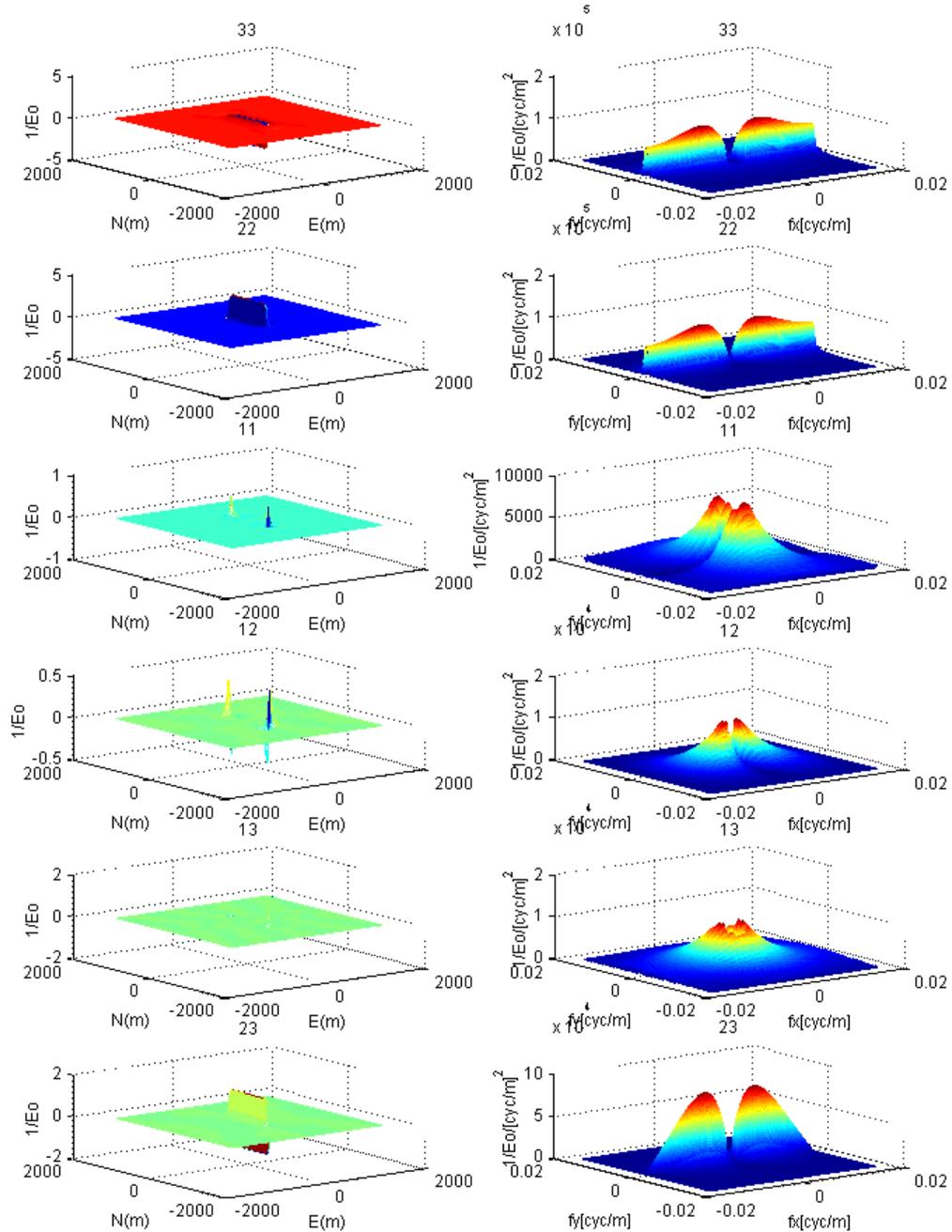


Figure 5.15: The matched filter functions and the response transfer functions (using Psd model 1) for the individual gradients due to the 2m x 2m x 1000m anomaly at the azimuthal orientation of 0° and 30m depth in the full geological background plus $1E^2/Hz$ white noise.

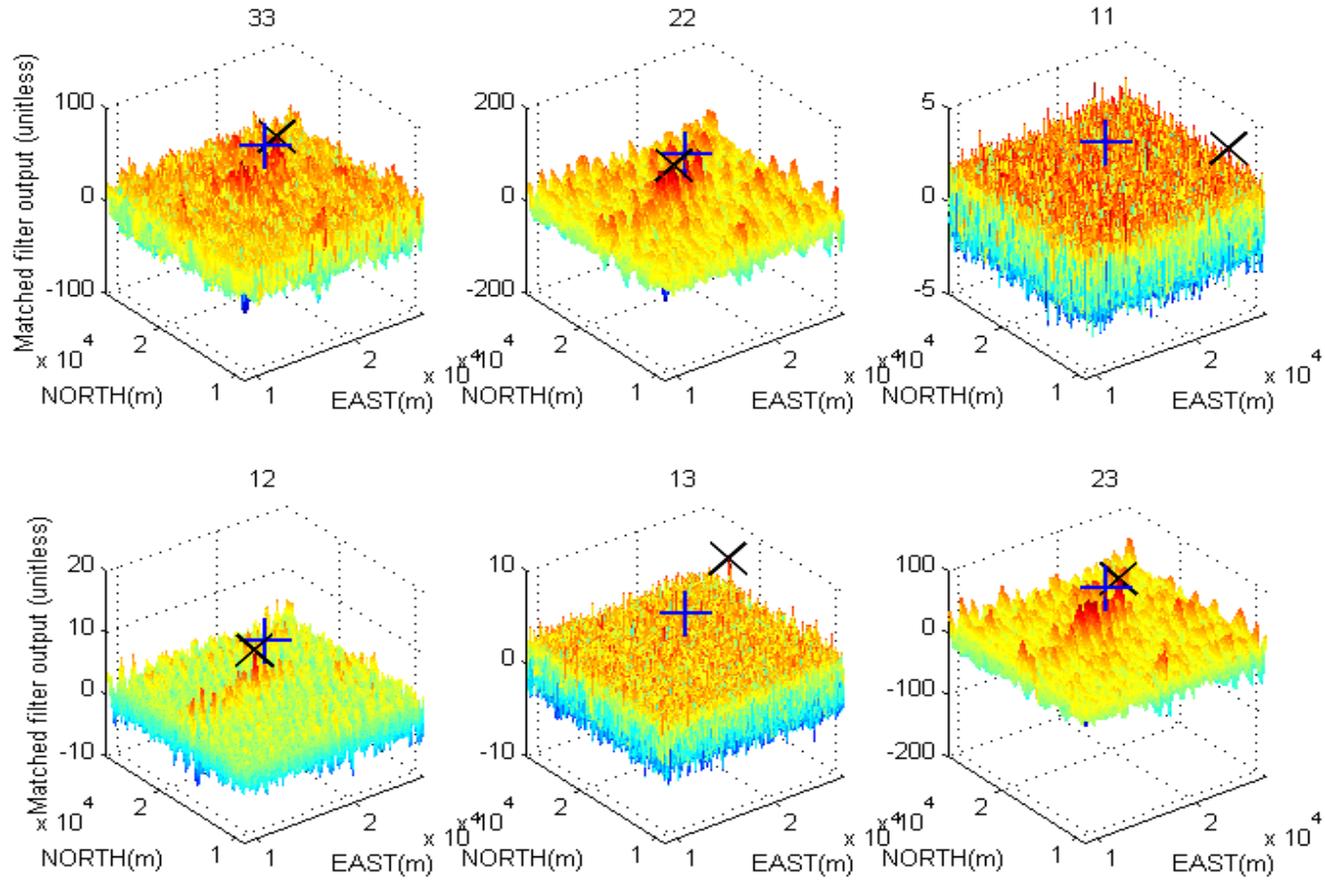


Figure 5.16: The matched filter outputs (using Psd model 1 for the full geological background plus $1E^2/\text{Hz}$ white noise); “x” indicates the location of highest peak; “+” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 0° and 30m depth.

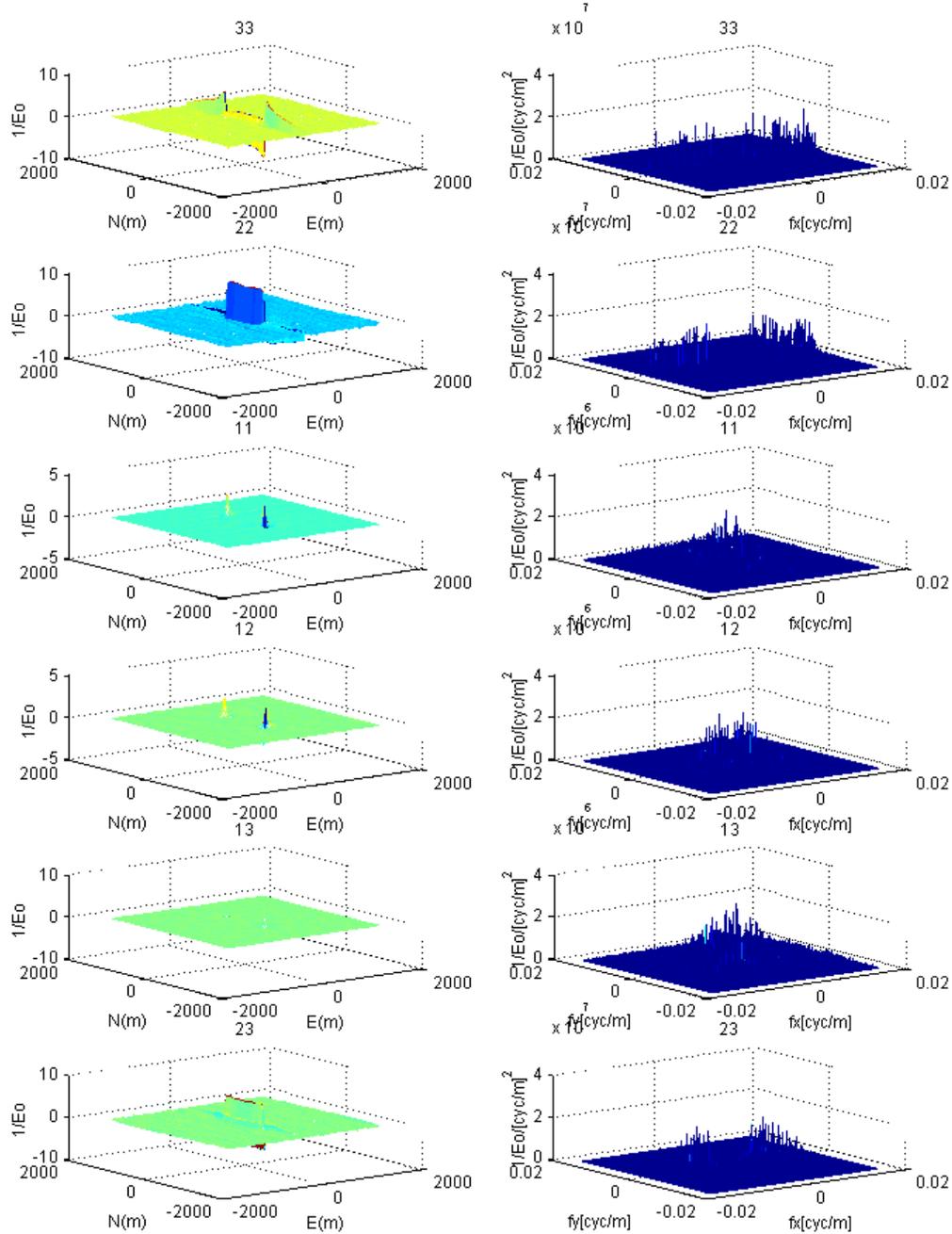


Figure 5.17: The matched filter functions and the transfer functions (using the empirical psd) for the individual gradients due to the 2m x 2m x 1000m anomaly at the azimuthal orientation of 0° and 30 m depth in the full geological background plus $1E^2/\text{Hz}$ white noise.

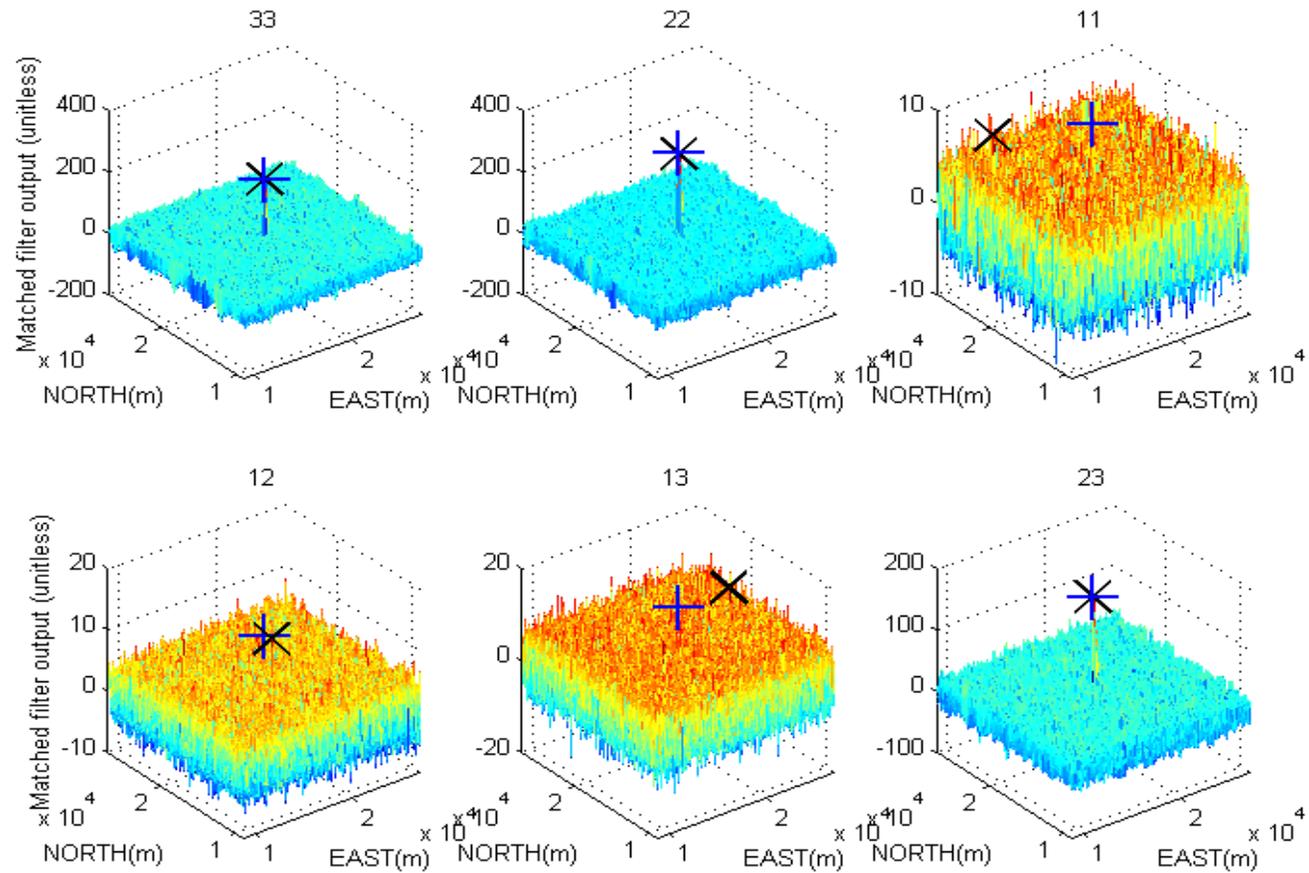


Figure 5.18: The matched filter outputs (using the empirical psd for the full geological background plus $1E^2/\text{Hz}$ white noise); “x” indicates the location of highest peak; “+” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 0° and 30m depth.

Type	Grad.	ΔN	ΔE	d^2	y_{\max}	POM = 0.001		POM = 0.010		POM = 0.050	
						η'	POF	η'	POF	η'	POF
a	Γ_{33}	540.0	1560.0	74.966	71.006	48.210	0.000	54.824	0.000	60.724	0.000
	Γ_{22}	-870.0	-1920.0	72.923	120.288	46.534	0.000	53.057	0.000	58.876	0.000
	Γ_{11}	-6870.0	6930.0	0.767	3.719	-1.939	0.987	-1.270	0.926	-0.674	0.779
	Γ_{12}	-840.0	-1770.0	0.824	10.262	-1.981	0.985	-1.288	0.922	-0.669	0.769
	Γ_{13}	5190.0	8250.0	1.562	6.407	-2.300	0.967	-1.345	0.859	-0.494	0.653
	Γ_{23}	570.0	1680.0	39.006	88.642	19.706	0.001	24.477	0.000	28.733	0.000
b	Γ_{33}	0.000	0.000	209.710	208.713	164.959	0.000	176.021	0.000	185.890	0.000
	Γ_{22}	0.000	0.000	291.702	291.370	238.923	0.000	251.969	0.000	263.609	0.000
	Γ_{11}	3420.0	-7560.0	5.080	9.561	-1.885	0.799	-0.163	0.529	1.373	0.271
	Γ_{12}	-690.0	180.0	4.819	10.459	-1.965	0.815	-0.288	0.552	1.208	0.291
	Γ_{13}	210.0	5280.0	9.843	13.591	0.148	0.481	2.544	0.209	4.682	0.068
	Γ_{23}	0.000	0.000	162.902	169.744	123.461	0.000	133.210	0.000	141.908	0.000

Table 5.7: Detectability and performance of the matched filter using (a) Psd model 1 and (b) the empirical psd in the full geological background plus $1E^2/\text{Hz}$ white noise; ΔN and ΔE are the difference of the location of the highest peak minus the true location; the 2m x 2m x 1000m anomaly at the azimuthal orientation of 0° and 30m depth.

5.4 CASE III: Anomaly detection in correlated noise using the combination of six gradient components

Rather than attempt to seek an accurate psd model for individual matched filters, as described in the previous case, this section shows how the matched filter using the combination of six gradient components with Psd model 1 performs in terms of signal detectability. The (vector) transfer function of the filter, $\tilde{\mathbf{H}}(f_1, f_2)$, is obtained from the equation (4.18) and is extended to the two dimensional case

$$\tilde{\mathbf{H}}(f_1, f_2) = \Delta x_1 \Delta x_2 \tilde{\Phi}_{MM}^{-1}(f_1, f_2) \tilde{\mathbf{S}}^*(f_1, f_2) \quad (5.10)$$

where the $M \times M$ cross psd matrix, $\tilde{\Phi}_{MM}$, between M sensors for $M \in \{1, 2, 3, 4, 5, 6\}$, is constructed according to (6.97) in Chapter 6 using (A.1) with the parameters of Psd model 1 plus the instrumental noise N_F in Table A.1: $\tilde{\Phi}_{MM} = \Delta x_1 \Delta x_2 \tilde{\Sigma}_{MM}$ (physical correlations between sensor outputs are considered) + N_F . The vectors $\tilde{\mathbf{H}}(f_1, f_2)$ and $\tilde{\mathbf{S}}(f_1, f_2)$ have size $M \times 1$. We also assume the matrix $\tilde{\Phi}_{MM}$ is invertible.

As shown in Figures 5.19 - 5.21 where the $2m \times 2m \times 1000m$ anomaly is oriented at $\beta' = 90^\circ$, 45° , and 0° , respectively, the filter outputs clearly indicate an improvement of signal detection and performance. The results are shown in Tables 5.8 – 5.10, where other combinations of gradients are also considered. For example, for $\beta' = 90^\circ$, the combination of Γ_{22} and Γ_{12} can not locate the true location where $\Delta N = -5940.0m$ and $\Delta E = 7770.0m$ (Table 5.8), which shows the low performance: POF = 0.970 for given POM = 0.001 and $d^2 = 1.461$. However, the highest peak occurs at the true location using all six components. The performance by the six combination is high: POF = 0.000 for given POM = 0.001 and $d^2 = 175.213$.

Another example is addressed to show how well the matched filter performs for combinations of gradients. Refer to CASE II in section 5.3 where Table 5.7a shows that no matched filter for individual gradients can detect the target anomaly for $\beta' = 0^\circ$, (e.g., the Γ_{33} matched filter mislocates the true location, i.e., $\Delta N = 540.0m$ and $\Delta E = 1560.0m$ whereas POF = 0.000 for given POM = 0.001 and $d^2 = 74.966$). In comparison, shown in Table 5.10, when all six combined gradients are used, the highest peak output occurs at the true location with a high performance. However, it should be noted that if POM is defined too high, the highest peak $y_{\max} = 142.656$ does not exceed the computed thresholds, i.e., $\eta' = 153.379$ and 144.360 , for given POM = 0.010 and 0.050, respectively. As such, we falsely reject H_0 . (Making a decision to either accept H_0 or H_1 is controlled by choosing POM) The POM = 0.001, where $\eta' = 134.250$ and POF = 0.000, is an appropriate level of significance for this case.

We further test the capability of the matched filter using the combination. But here, consider the anomaly for $\beta' = 90^\circ$ at the deeper depth = 60m where the gradient signals are much smaller. As expected, the matched filter using all six combined gradients yields the most reliable detection, shown in Figure 5.22. The highest peak output exists at the true location with POF = 0.254 for POM = 0.001 and $d^2 = 14.068$,

shown in Table 5.11. However, the performance of the combination filter is downgraded as compared to the case where $\beta' = 90^\circ$ and depth = 30m: POF = 0.000 for POM = 0.001 and $d^2 = 175.213$ (Table 5.8). Overall, the matched filters using all six gradients with Psd model 1 are successful in locating the true anomaly and have an increased performance.

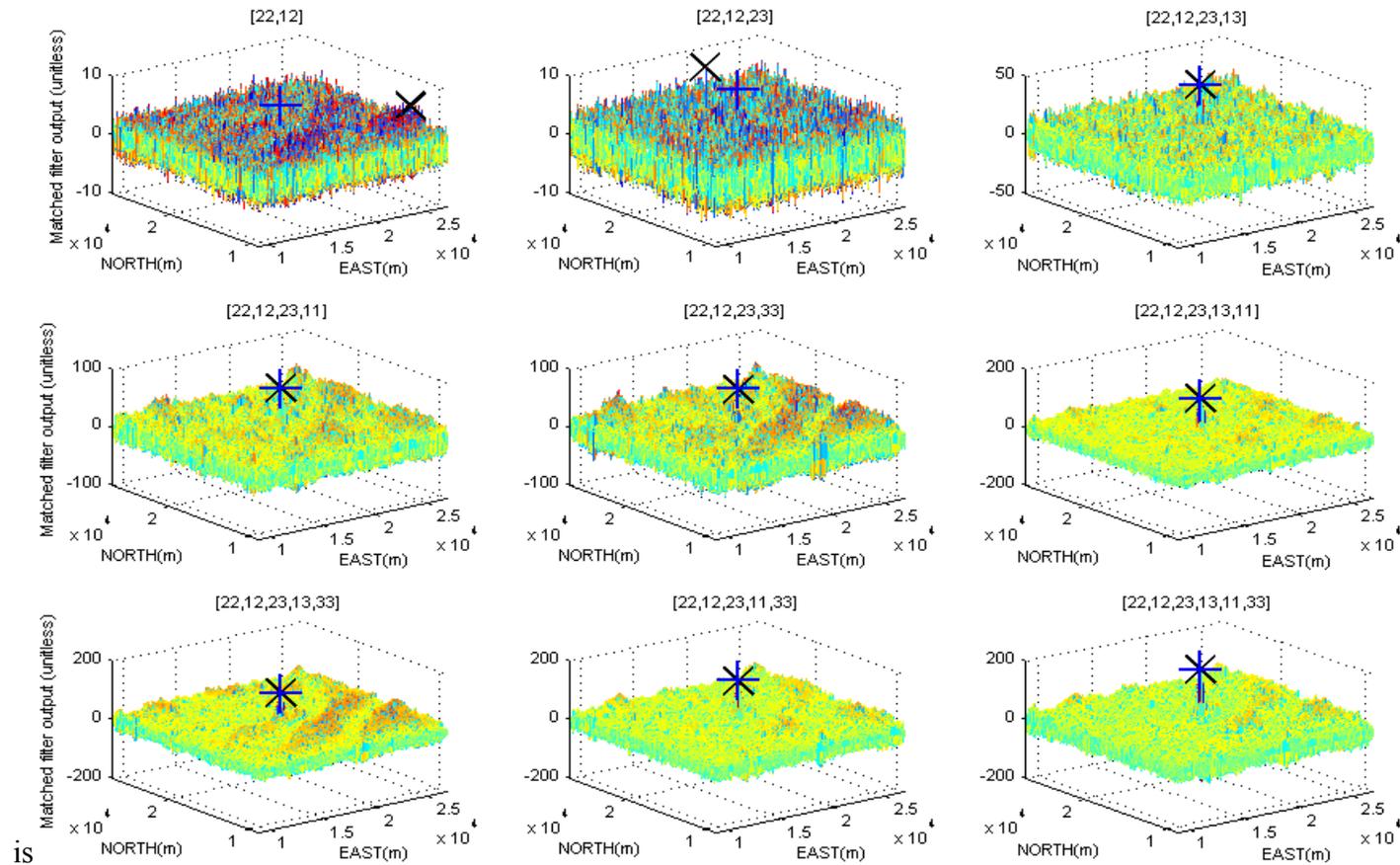


Figure 5.19: The matched filter output of the combination of six gradients in the full geological background plus $1E^2/\text{Hz}$ white noise using Psd model 1; “x” indicates the location of highest peak; “+” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 90° and 30m depth.

Orient.	Grad.	ΔN (m)	ΔE (m)	d^2	y_{\max}	POM = 0.001		POM = 0.010		POM = 0.050	
						η'	POF	η'	POF	η'	POF
90°	2,4	-5940.0	7770.0	1.461	6.052	-2.274	0.970	-1.351	0.868	-0.527	0.669
	2,4,6	6120.0	1530.0	2.893	9.019	-2.363	0.918	-1.064	0.734	0.095	0.478
	2,4,6,5	0.0	0.0	38.474	48.312	19.306	0.001	24.044	0.000	28.271	0.000
	2,4,6,3	0.0	0.0	71.403	79.042	45.291	0.000	51.745	0.000	57.504	0.000
	2,4,6,1	0.0	0.0	73.694	76.954	47.166	0.000	53.724	0.000	59.574	0.000
	2,4,6,5,3	0.0	0.0	106.049	121.491	74.225	0.000	82.092	0.000	89.110	0.000
	2,4,6,5,1	0.0	0.0	108.142	115.451	76.006	0.000	83.950	0.000	91.037	0.000
	2,4,6,3,1	0.0	0.0	141.005	155.654	104.310	0.000	113.381	0.000	121.473	0.000
	2,4,6,5,3,1	0.0	0.0	175.213	192.861	134.308	0.000	144.420	0.000	153.440	0.000

NOTE:
 $1 = \Gamma_{33}$, $2 = \Gamma_{22}$, $3 = \Gamma_{11}$, $4 = \Gamma_{12}$, $5 = \Gamma_{13}$, $6 = \Gamma_{23}$

Table 5.8: Detectability and performance of the combination of matched filter using Psd model 1 in the full geological background plus $1E^2/\text{Hz}$ white noise; ΔN and ΔE are the difference of the location of the highest peak minus the true location of the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90° and 30m depth.

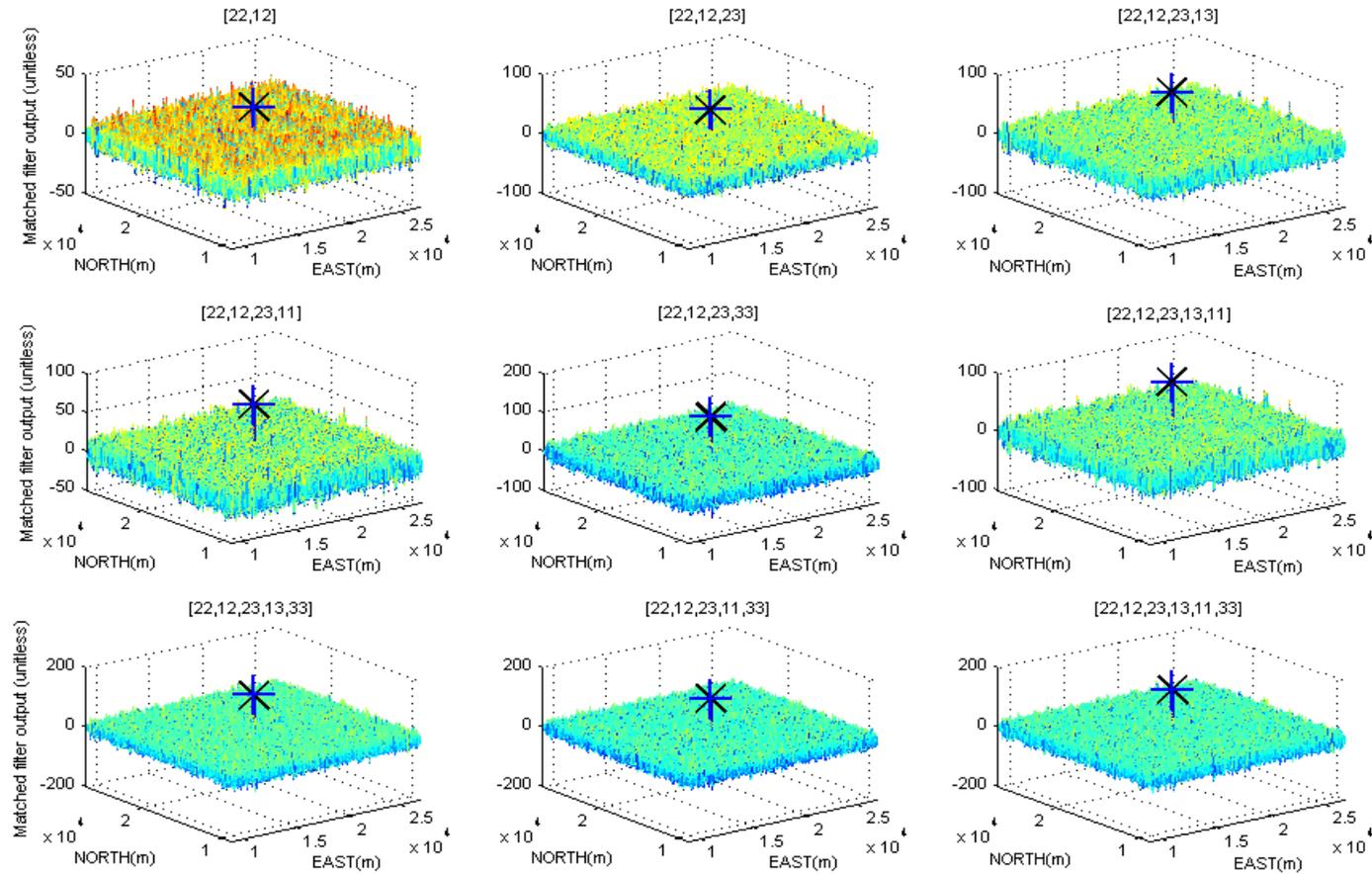


Figure 5.20: The matched filter output of the combination of six gradients in the full geological background plus $1E^2/\text{Hz}$ white noise using PSD model 1; “ \times ” indicates the location of highest peak; “ $+$ ” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 45° and 30m depth.

Orient.	Grad.	ΔN (m)	ΔE (m)	d^2	y_{\max}	POM = 0.001		POM = 0.010		POM = 0.050	
						η'	POF	η'	POF	η'	POF
45°	2,4	30.0	30.0	30.994	28.545	13.790	0.007	18.042	0.001	21.836	0.000
	2,4,6	-60.0	-60.0	57.614	52.813	34.158	0.000	39.956	0.000	45.129	0.000
	2,4,6,5	-60.0	-60.0	84.070	82.164	55.736	0.000	62.740	0.000	68.988	0.000
	2,4,6,3	0.0	0.0	73.372	68.446	46.901	0.000	53.445	0.000	59.282	0.000
	2,4,6,1	0.0	0.0	117.581	103.604	84.072	0.000	92.356	0.000	99.745	0.000
	2,4,6,5,3	-60.0	-60.0	99.720	96.342	68.861	0.000	76.490	0.000	83.295	0.000
	2,4,6,5,1	0.0	0.0	143.700	130.771	106.656	0.000	115.813	0.000	123.982	0.000
	2,4,6,3,1	0.0	0.0	133.139	119.166	97.482	0.000	106.296	0.000	114.159	0.000
	2,4,6,5,3,1	0.0	0.0	159.194	146.547	120.204	0.000	129.842	0.000	138.441	0.000

NOTE:
1 = Γ_{33} , 2 = Γ_{22} , 3 = Γ_{11} , 4 = Γ_{12} , 5 = Γ_{13} , 6 = Γ_{23}

Table 5.9: Detectability and performance of the combination of matched filter using Psd model 1 in the full geological background plus $1E^2/\text{Hz}$ white noise; ΔN and ΔE are the difference of the location of the highest peak minus the true location of the 2m x 2m x 1000m anomaly at the azimuthal orientation of 45° and 30m depth.

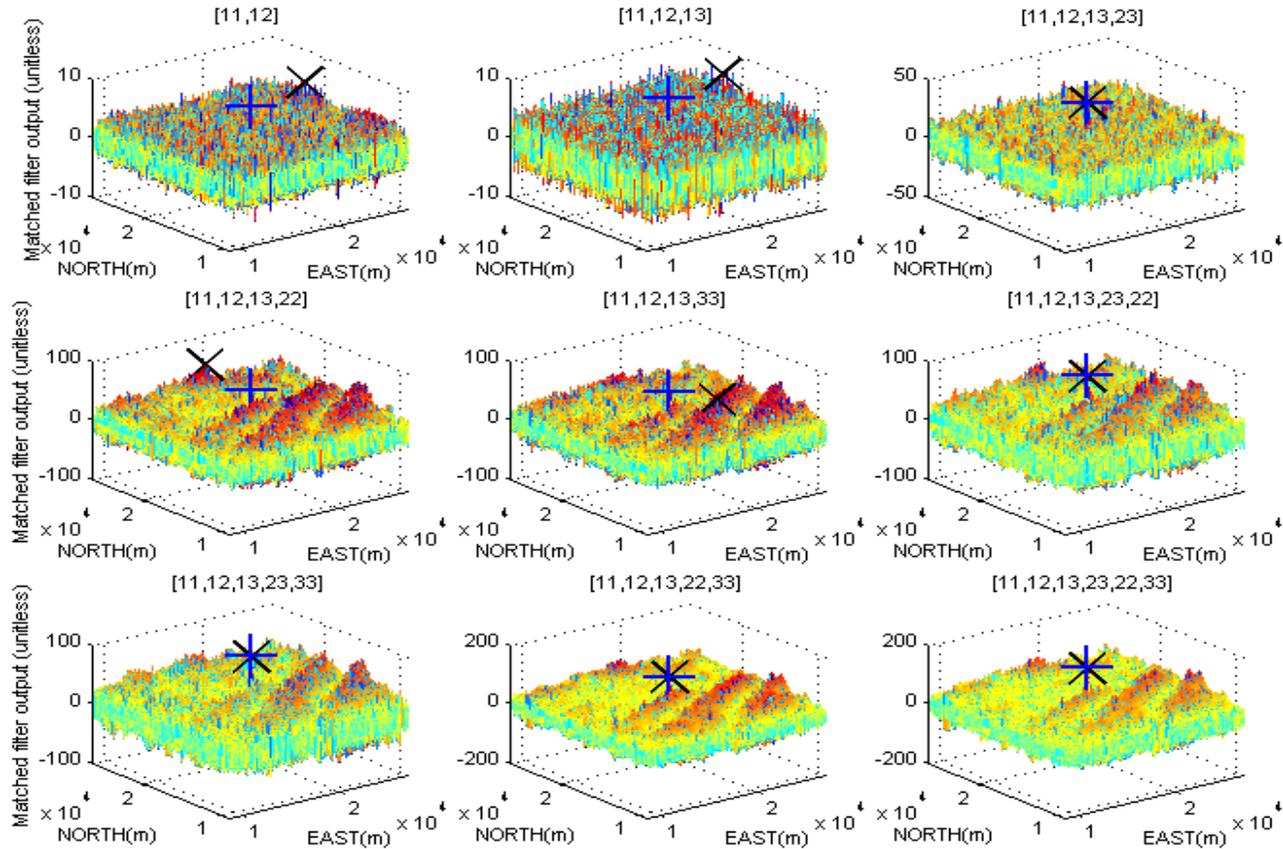


Figure 5.21: The matched filter output of the combination of six gradients in the full geological background plus $1E^2/\text{Hz}$ white noise using Psd model 1; “x” indicates the location of highest peak; “+” indicates the true location of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at the azimuthal orientation of 0° and 30m depth.

Orient.	Grad.	ΔN (m)	ΔE (m)	d^2	y_{\max}	POM = 0.001		POM = 0.010		POM = 0.050	
						η'	POF	η'	POF	η'	POF
0°	3,4	2490.0	7290.0	1.461	6.296	-2.274	0.970	-1.351	0.868	-0.527	0.669
	3,4,5	-3030.0	9000.0	2.893	8.045	-2.363	0.918	-1.064	0.734	0.095	0.478
	3,4,5,6	0.0	0.0	38.471	34.559	19.304	0.001	24.042	0.000	28.269	0.000
	3,4,5,2	7560.0	1110.0	71.403	60.961	45.290	0.000	51.745	0.000	57.504	0.000
	3,4,5,1	-3960.0	1800.0	73.690	59.023	47.163	0.000	53.720	0.000	59.570	0.000
	3,4,5,6,2	0.0	0.0	106.014	85.513	74.196	0.000	82.062	0.000	89.079	0.000
	3,4,5,6,1	0.0	0.0	108.251	91.010	76.099	0.000	84.047	0.000	91.137	0.000
	3,4,5,2,1	0.0	0.0	141.007	110.614	104.312	0.000	113.383	0.000	121.475	0.000
	3,4,5,6,2,1	0.0	0.0	175.148	142.656	134.250	0.000	144.360	0.000	153.379	0.000

NOTE: 1 = Γ_{33} , 2 = Γ_{22} , 3 = Γ_{11} , 4 = Γ_{12} , 5 = Γ_{13} , 6 = Γ_{23}

Table 5.10: Detection and performance of the combination of matched filter using Psd model 1 in the full geological background plus $1E^2/\text{Hz}$ white noise; ΔN and ΔE are the difference of the location of the highest peak minus the true location of the 2m x 2m x 1000m anomaly at the azimuthal orientation of 0° and 30m depth.

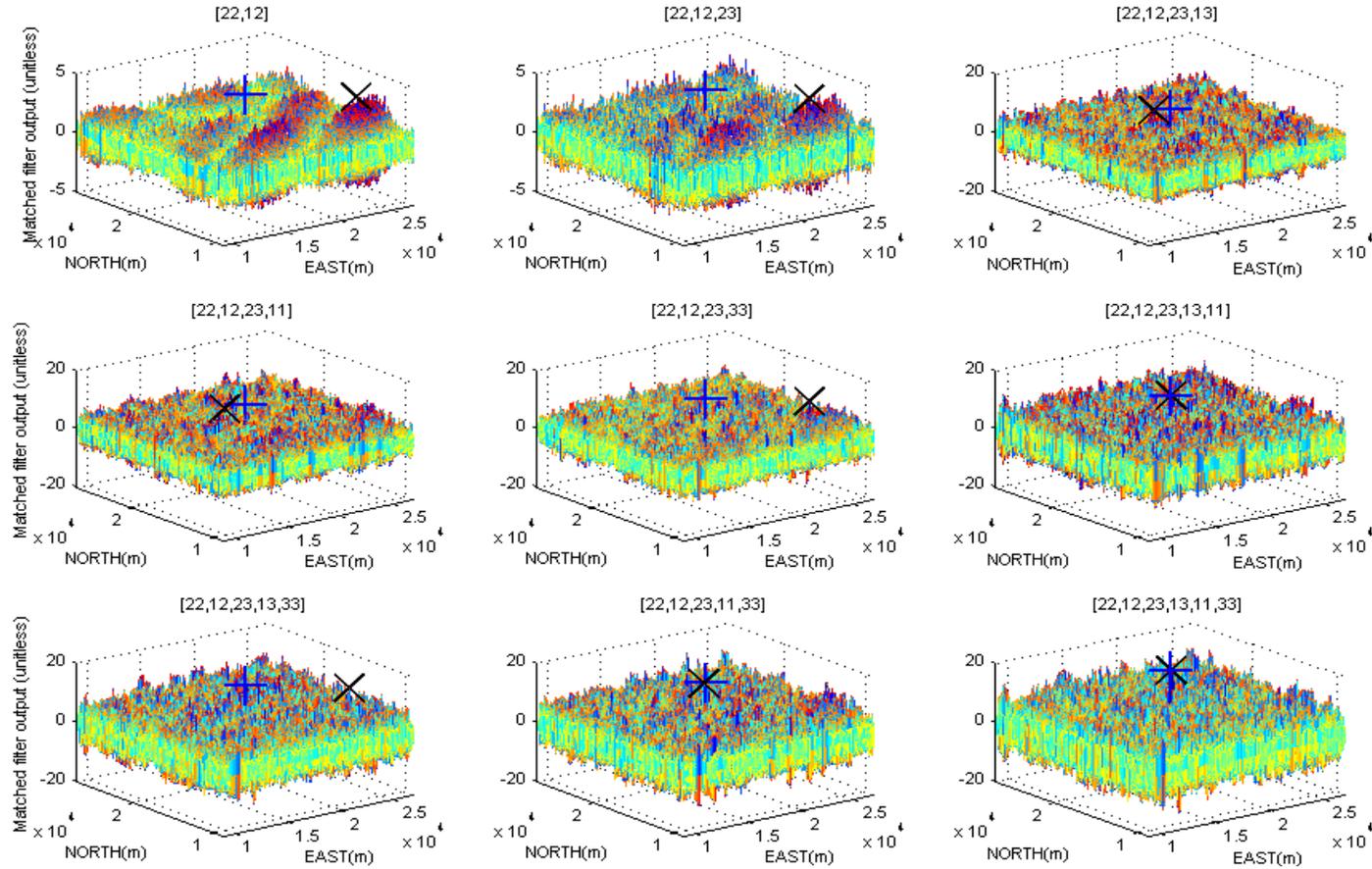


Figure 5.22: The matched filter output of the combination of six gradients in the full geological background plus $1E^2/\text{Hz}$ white noise using Psd model 1; “x” indicates the location of highest peak; “+” indicates the true location of the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90° and 60m depth.

Orient.	Grad.	ΔN (m)	ΔE (m)	d^2	y_{\max}	POM = 0.001		POM = 0.010		POM = 0.050	
						η'	POF	η'	POF	η'	POF
90°	2,4	-5760.0	6030.0	0.237	3.852	-1.267	0.995	-0.895	0.967	-0.564	0.877
	2,4,6	-6600.0	4710.0	0.442	4.183	-1.612	0.992	-1.105	0.952	-0.652	0.836
	2,4,6,5	480.0	-1230.0	5.220	10.540	-1.840	0.790	-0.095	0.517	1.462	0.261
	2,4,6,3	-120.0	-2130.0	5.190	10.606	-1.850	0.792	-0.110	0.519	1.442	0.263
	2,4,6,1	-5250.0	5700.0	5.457	12.298	-1.762	0.775	0.022	0.496	1.614	0.245
	2,4,6,5,3	0.000	0.000	9.600	13.674	0.025	0.497	2.392	0.220	4.504	0.073
	2,4,6,5,1	-5250.0	5700.0	9.855	14.957	0.154	0.480	2.552	0.208	4.692	0.067
	2,4,6,3,1	-90.0	-240.0	9.831	15.568	0.142	0.482	2.537	0.209	4.674	0.068
	2,4,6,5,3,1	0.000	0.000	14.068	19.750	2.478	0.254	5.343	0.077	7.899	0.018

NOTE: 1 = Γ_{33} , 2 = Γ_{22} , 3 = Γ_{11} , 4 = Γ_{12} , 5 = Γ_{13} , 6 = Γ_{23}

Table 5.11: Detection and performance of the combination of matched filter using Psd model 1 in the full geological background plus $1E^2/\text{Hz}$ white noise; ΔN and ΔE are the difference of the location of the highest peak minus the true location of the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90° and 60m depth.

5.5 CASE IV: Anomaly detectability versus sizes and depths

Thus far, we have discussed the anomaly detection by matched filtering techniques in $1E^2/\text{Hz}$ white noise and full geological background plus $1E^2/\text{Hz}$ white noise. The $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly was considered. The numerical results indicate that, overall, the Γ_{33} matched filter is able to detect the anomaly best among other components. Psd modeling affects the filter detectability, as compared to the case where the empirical psd's were applied. Although more accurate psd models are needed to enhance the capability of matched filters, the matched filter using all six gradients with Psd model 1 shows an improvement of signal detection and performance superior to the individual matched filters with Psd model 1.

This section examines the detectability associated with sizes and depths of anomalies at the azimuthal orientations of 90° , 45° , and 0° . More types of noise backgrounds are included to see how they affect the detectability associated with these sizes and depths using both psd models and empirical psd's. The noise backgrounds are as follows (i) full geological background plus $1E^2/\text{Hz}$ white noise (or instrumental white noise), (ii) full geological background plus $1E^2/\text{Hz}$ white noise minus a geological model obtained as described in section 5.1 (see Footnote 6), (iii) the longer than 5km wavelength GGSS data plus $1E^2/\text{Hz}$ white noise, and (iv) only $1E^2/\text{Hz}$ white noise (considered as the best-case detection situation).

The parameters of anomalies for the investigation are defined in Table 5.12. Figure 5.23 shows the test strategy to determine anomaly detectability with respect to sizes and depths. To determine the maximum depth, D_{\max} , at which different anomalies can be detected using individual gradients and all six combined gradients, the detectability criterion is defined by three steps as follows:

- I. At depth i for $i \in \{10\text{m}, 20\text{m}, 30\text{m}, \dots, D_{\max}\}$, select the location, (N_i, E_i) , of the matched filter output where the highest peak $y_{\max,i}$ occurs.
- II. If $|N_i - N_0| \leq \delta x_1$ and $|E_i - E_0| \leq \delta x_2$, then the target signal is detected at depth i .
- III. Return to step I if the anomaly is located at depth i ; otherwise declare depth $i-1$ as the maximum depth D_{\max} .

Type	Width (m)	Height (m)	Length (m)
I	5	5	5,10, and 100
II	2	2	10,100, and 1000

Table 5.12: The various types of target anomalies.

The symbol (N_0, E_0) is the true location of target anomalies of types I and II (which is given). Based on a number of numerical experiments, the tolerances of $\delta x_1 = 90\text{m}$ and $\delta x_2 = 90\text{m}$ are an appropriate choice for the criterion of step II. The values in Tables 5.13 to 5.22 provide a detailed analysis, which we briefly discuss in terms of how the factors stated above affect the anomaly detectability.

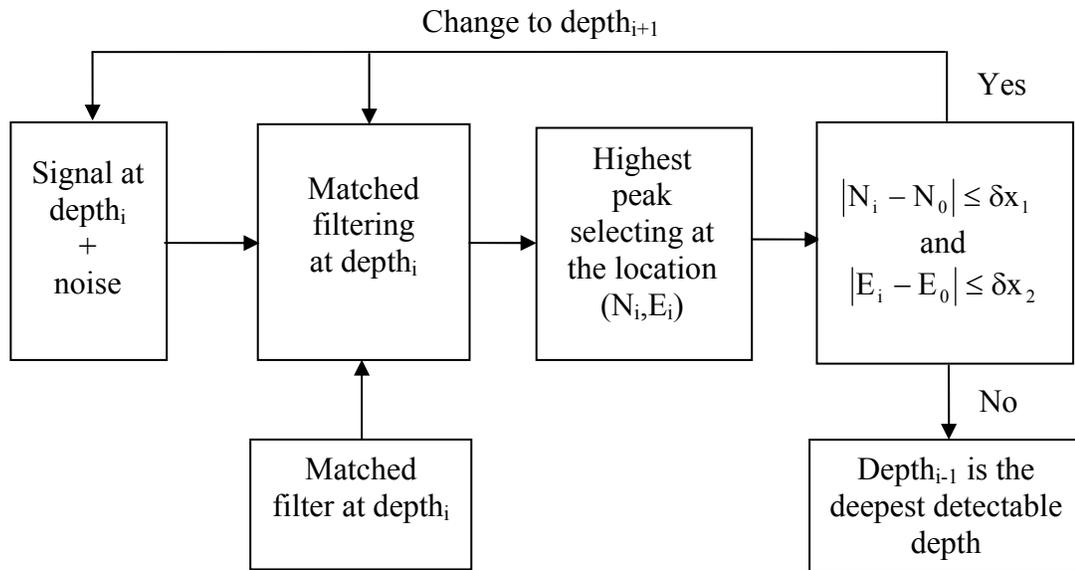


Figure 5.23: Block diagram of anomaly detectability criteria.

The maximum detectable depths obtained from the matched filters for individual gradients using the psd models are summarized in Tables 5.13 – 5.15 at $\beta' = 90^\circ, 45^\circ,$ and 0° , respectively. These values of the detectable depths are compared to the results for the case at which the empirical psd's are applied for the detection, summarized in Tables 5.16 – 5.18. The results show that the anomalies are detectable at deeper depths using the filters with the empirical psd's. We also see that the size of anomaly affects the detectability, e.g., in Table 5.13, the matched filters can detect the $5\text{m} \times 5\text{m} \times 10\text{m}$ as deep as 20m , as compared to the deepest depth of 50m for the $5\text{m} \times 5\text{m} \times 100\text{m}$ anomaly in full geological background plus $1\text{E}^2/\text{Hz}$ white noise. However, the background noise only slightly affects the detectability by using the matched filters with the psd models, e.g., see the detectable depths of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly in Tables 5.13.- 5.15 hardly changes with different background noises. (However, when wrong psd models are applied to the matched filters, the anomaly can not be detected at all; not shown here) By comparing the detection results, the orientation also affects the detectability, particularly

Γ_{12} whose magnitude increases at $\beta' = 45^\circ$. (More details related to orientations will be discussed in the next section).

As discussed in CASE II, the psd modeling affects the detectability. The numerical details of model versus empirical psd are summarized in Tables 5.16 – 5.18. For example, the anomaly having the cross section area of 2m x 2m can not be detected at all using the Γ_{22} matched filters using the psd models in Table 5.13 whereas it is detectable at depth of 10m using the empirical psd (Table 5.16). Overall, the Γ_{33} matched filters are able to detect the target anomalies best among other component filters. Also we found from our numerical results that the detectability of the Γ_{33} matched filter does not depend on the orientations $\beta' = 90^\circ, 45^\circ$, and 0° since Γ_{33} is independent of orientation according to eq. (B.5).

The matched filters using all six combined gradients improve the anomaly detectability. (For simplicity, assume no correlation between sensors in the white noise environment) For instance, consider the 2m x 2m x 1000m anomaly at $\beta' = 90^\circ$ in $1E^2/\text{Hz}$ white noise. The deepest depth of 80m can be detected using the combination (Table 5.19), compared with the deepest depth = 60m in Table 5.13 where the Γ_{13} matched filter performs best to detect the same anomaly (Psd model 4 is used in both cases).

Table 5.20 shows that the location of the target affects the detectability. This is due to random noise effects (from $1E^2/\text{Hz}$ white noise) and geological interferences (from full geological background) at that location. Here, we only show the cases of the target anomalies at $\beta' = 90^\circ$. Note that using the same noise contents, e.g., $1E^2/\text{Hz}$ white noise, the detectable depths, D_{\max} , may not be equal at different true locations, e.g. $D_{\max} = 80\text{m}$ at the true location of (N24000m, E24000m) versus $D_{\max} = 100\text{m}$ at the true location of (N24000m, E12000m).

We also consider when the true location of the target anomaly is inside a grid square. Table 5.21 shows the detectability results. They show that the detectability is slightly downgraded when the true location is not on a grid intersection. For instance, consider the 2m x 2m x 100m anomaly having the true location at N18013m and E18018m. In the case of full geological background plus $1E^2/\text{Hz}$ white noise, the Γ_{33} filter using the empirical psd gives $D_{\max} = 30\text{m}$, whereas the Γ_{33} filter in Table 5.16, where the true location is at (N18000m, E18000m), yields $D_{\max} = 40\text{m}$.

Finally, the detectability also depends on the resolution of a gradiometric survey. The numerical analyses in Table 5.22 indicate the effects of the spatial interval on the detectability. For instance, compare the cross-track spacing $\Delta x_1 = 30\text{m}$ and the along track spacing $\Delta x_2 = 15\text{m}$ with $\Delta x_1 = 15$ and $\Delta x_2 = 5\text{m}$ for detecting the 2m x 2m x 100m anomaly. The latter case yields $D_{\max} = 60\text{m}$ whereas $D_{\max} = 40\text{m}$ for the former.

Anomaly		Maximum depth (m)					
Width x Height (m ²)	Length(m)	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
Psd model 1 for geological background + 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	50	10	50	30	50	20
2x2	10	10	-	10	-	-	-
	100	20	-	20	-	-	-
	1000	30	-	40	-	50	-
Psd model 2 for geological background +1E ² /Hz white noise - geological model							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	70	20	60	30	40	20
2x2	10	10	-	10	-	-	-
	100	20	-	20	-	-	-
	1000	30	-	40	-	40	-
Psd model 3 for the longer-5km-wavelength GGSS data + 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	70	30	60	40	50	40
2x2	10	10	-	10	-	-	-
	100	20	-	20	-	-	-
	1000	40	-	40	-	60	-
Psd model 4 for 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	70	30	60	40	50	40
2x2	10	10	-	10	-	-	-
	100	20	-	20	-	-	-
	1000	40	-	40	-	60	-

Table 5.13: Anomaly detectability versus sizes and depths at the azimuthal orientation of 90° by the matched filters using the psd models for various noise environments; the center of anomaly is at N18000m and E18000m.

Anomaly		Maximum depth (m)					
Width x Height (m ²)	Length(m)	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
Psd model 1 for geological background + 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	50	20	60	40	70	30
2x2	10	10	-	10	-	-	-
	100	20	10	10	10	10	-
	1000	30	10	20	30	30	10
Psd model 2 for geological background +1E ² /Hz white noise - geological model							
5x5	5	20	10	10	-	-	-
	10	20	10	20	10	-	-
	100	60	30	50	30	40	30
2x2	10	10	-	10	-	-	-
	100	20	10	10	10	-	-
	1000	40	20	20	30	30	20
Psd model 3 for the longer-5km-wavelength GGSS data + 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	70	30	60	30	50	40
2x2	10	10	-	10	-	-	-
	100	20	10	10	10	10	-
	1000	30	20	30	40	30	20
Psd model 4 for 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	70	30	60	30	50	40
2x2	10	10	-	10	-	-	-
	100	20	10	10	10	10	-
	1000	30	20	30	40	30	20

Table 5.14: Anomaly detectability versus sizes and depths at the azimuthal orientation of 45° by the matched filters using the psd models for various noise environments; the center of anomaly is at N18000m and E18000m.

Anomaly		Maximum depth (m)					
Width x Height (m ²)	Length(m)	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
Psd model 1 for geological background + 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	40	30	40	30	40	30
2x2	10	10	10	-	-	-	-
	100	20	10	-	-	-	-
	1000	20	20	-	-	-	-
Psd model 2 for geological background + 1E ² /Hz white noise - geological model							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	60	40	40	30	40	30
2x2	10	10	10	-	-	-	-
	100	20	10	-	-	-	-
	1000	30	30	-	-	-	10
Psd model 3 for the longer-5km-wavelength GGSS data + 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	70	40	50	40	40	50
2x2	10	10	10	-	-	-	-
	100	20	10	-	-	-	-
	1000	30	40	-	-	-	30
Psd model 4 for 1E ² /Hz white noise							
5x5	5	20	10	10	-	-	-
	10	20	10	20	-	-	-
	100	70	40	50	40	40	50
2x2	10	10	10	-	-	-	-
	100	20	10	-	-	-	-
	1000	30	40	-	-	-	30

Table 5.15: Anomaly detectability versus sizes and depths at the azimuthal orientation of 0° by the matched filters using the psd models for various noise environments; the center of anomaly is at N18000m and E18000m.

Anomaly		Maximum depth (m)					
Width x Height (m ²)	Length(m)	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
The empirical psd for geological background + 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	30	20	30	10	20	30
	100	100	60	70	50	80	70
2x2	10	20	10	10	-	-	-
	100	40	10	30	10	30	20
	1000	50	10	60	-	60	20
The empirical psd for geological background + 1E ² /Hz white noise - geological model							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	110	50	90	50	80	80
2x2	10	20	10	10	-	-	-
	100	40	10	30	10	20	20
	1000	60	20	90	-	70	20
The empirical psd for the longer-5km-wavelength GGSS data + 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	110	50	80	60	90	70
2x2	10	20	10	10	-	-	-
	100	40	10	30	10	20	20
	1000	60	20	70	-	90	20
The empirical psd for 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	120	50	80	50	80	70
2x2	10	20	10	10	-	-	-
	100	40	10	30	10	30	20
	1000	70	20	70	-	80	20

Table 5.16: Anomaly detectability versus sizes and depths at the azimuthal orientation of 90° by the matched filters using the empirical psd's for various noise environments; the center of anomaly is at N18000m and E18000m.

Anomaly		Maximum depth (m)					
Width x Height (m ²)	Length(m)	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
The empirical psd for geological background + 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	30	20	30	10	20	30
	100	80	60	80	70	70	100
2x2	10	20	10	10	-	-	-
	100	30	20	20	20	30	30
	1000	80	40	40	60	50	60
The empirical psd for geological background + 1E ² /Hz white noise - geological model							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	110	60	80	60	70	90
2x2	10	20	10	10	-	-	-
	100	40	20	20	20	30	30
	1000	110	40	50	60	50	80
The empirical psd for the longer-5km-wavelength GGSS data + 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	100	60	80	60	70	80
2x2	10	20	10	10	-	-	-
	100	40	20	20	20	30	20
	1000	100	40	50	60	50	70
The empirical psd for 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	100	60	80	60	70	80
2x2	10	20	10	10	-	-	-
	100	40	20	20	20	30	20
	1000	100	40	50	60	50	70

Table 5.17: Anomaly detectability versus sizes and depths at the azimuthal orientation of 45° filters using the empirical psd's for various noise environments; the center of anomaly is at N18000m and E18000m.

Anomaly		Maximum depth (m)					
Width x Height (m ²)	Length(m)	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
The empirical psd for geological background + 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	30	20	30	10	20	30
	100	100	60	70	50	80	80
2x2	10	20	10	10	-	-	-
	100	20	10	10	-	-	-
	1000	50	120	10	-	20	60
The empirical psd for geological background + 1E ² /Hz white noise - geological model							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	100	60	80	50	80	80
2x2	10	20	10	10	-	-	-
	100	40	20	10	-	20	40
	1000	70	60	10	-	50	80
The empirical psd for the longer-5km-wavelength GGSS data + 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	110	60	80	60	90	70
2x2	10	20	10	10	-	-	-
	100	40	20	10	-	20	30
	1000	60	70	10	-	40	70
The empirical psd for 1E ² /Hz white noise							
5x5	5	30	20	20	-	-	-
	10	40	20	30	10	20	30
	100	110	60	70	50	80	80
2x2	10	20	10	10	-	-	-
	100	40	20	10	-	20	40
	1000	70	70	10	-	90	80

Table 5.18: Anomaly detectability versus sizes and depths at the azimuthal orientation of 0 ° by the matched filters using the empirical psd's for various noise environments; the center of anomaly is at N18000m and E18000m.

Anomaly		Geo. Back. + 1E ² /Hz white noise	Geo. Back. + 1E ² /Hz white noise – Geo. model	The longer-5km- wavelength GGSS data + 1E ² /Hz white noise	1E ² /Hz white noise
Orient.	Width x Height x Length (m ³)	Depth (m)	Depth (m)	Depth (m)	Depth (m)
90°	5x5x5	20	20	20	20
	5x5x10	20	20	30	30
	5x5x100	70	70	90	80
	2x2x10	10	10	10	10
	2x2x100	30	40	30	30
	2x2x1000	60	70	80	80
45°	5x5x5	20	20	20	20
	5x5x10	20	20	30	30
	5x5x100	80	60	80	80
	2x2x10	10	10	10	10
	2x2x100	20	20	20	20
	2x2x1000	50	50	60	60
0°	5x5x5	20	20	20	20
	5x5x10	20	20	30	30
	5x5x100	60	60	80	80
	2x2x10	10	10	10	10
	2x2x100	20	20	20	20
	2x2x1000	40	40	50	50

Table 5.19: Anomaly detectability versus sizes, depths, and orientations by the matched filters for all six combined gradients using psd models; for 1E²/Hz white noise, assume no correlation between sensors; the center of anomaly is at N18000m and E18000m.

Anomaly		Maximum depth (m)					
Noise background.	The center' location	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
Geo. Back. + 1E²/Hz white noise	(N18000m, E18000m)	50	10	60	-	60	20
	(N24000m, E24000m)	70	10	60	-	50	20
	(N24000m, E12000m)	60	10	90	20	60	10
	(N12000m, E12000m)	50	10	50	-	60	20
	(N12000m, E24000m)	50	10	60	-	50	10
Geo. Back. + 1E²/Hz white noise – Geo. model	(N18000m, E18000m)	60	20	90	-	70	20
	(N24000m, E24000m)	90	10	60	-	50	20
	(N24000m, E12000m)	70	10	100	-	30	10
	(N12000m, E12000m)	70	10	60	-	60	20
	(N12000m, E24000m)	60	10	90	-	60	10
1E²/Hz white noise	(N18000m, E18000m)	70	20	70	-	80	20
	(N24000m, E24000m)	80	10	50	-	50	20
	(N24000m, E12000m)	100	10	110	10	50	10
	(N12000m, E12000m)	100	10	90	-	70	30
	(N12000m, E24000m)	100	20	80	-	60	10

Table 5.20: Anomaly detectability of the matched filters using the empirical psd's versus locations and depths in various noise environments; the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90°.

Anomaly	Geo. Back. + 1E ² /Hz white noise			Geo. Back. + 1E ² /Hz white noise – Geo. model			The longer-5km-wavelength GGSS data + 1E ² /Hz white noise			1E ² /Hz white noise		
	Width x Height x Length (m ³)	Depth (m)	ΔN (m)	ΔE (m)	Depth (m)	ΔN (m)	ΔE (m)	Depth (m)	ΔN (m)	ΔE (m)	Depth (m)	ΔN (m)
2x2x10	10	-13	+12	10	-13	+12	10	-13	+12	10	-13	+12
	10	-13	+12	10	-13	+12	10	-13	+12	10	-13	+12
	<i>10</i>	<i>-13</i>	<i>+12</i>	<i>10</i>	<i>-13</i>	<i>+12</i>	<i>10</i>	<i>-13</i>	<i>+12</i>	<i>10</i>	<i>-13</i>	<i>+12</i>
2x2x100	10	-13	+12	10	-13	+12	10	-13	+12	10	-13	+12
	30	-13	+12	30	-13	+12	40	-13	-18	40	-13	-18
	<i>10</i>	<i>-13</i>	<i>+12</i>	<i>20</i>	<i>-13</i>	<i>-18</i>	<i>20</i>	<i>-13</i>	<i>-18</i>	<i>20</i>	<i>-13</i>	<i>-18</i>
2x2x1000	20	-13	-18	20	-13	-18	20	-13	-18	20	-13	-18
	40	-13	+12	60	-13	-18	70	-13	-18	70	-13	-18
	<i>50</i>	<i>-13</i>	<i>-18</i>	<i>60</i>	<i>-13</i>	<i>-18</i>	<i>70</i>	<i>-13</i>	<i>-78</i>	<i>70</i>	<i>-13</i>	<i>-78</i>

NOTE:
Normal fonts indicate the detection results by the Γ_{33} matched filter with the “smooth” psd models.
Bold fonts indicate the detection results by the Γ_{33} matched filter with the empirical psd’s.
Italic fonts indicate the detection results by the combination of the matched filters for all six combined gradients using the psd models (assume no correlation between sensors).

Table 5.21: The anomaly detectability of the Γ_{33} matched filter using the psd models versus the empirical psd’s for various noise backgrounds. The matched filter for all six combined gradients are used in comparisons. The center of the anomaly at the azimuthal orientation of 90° is located at N=18013m and E=18018m.

Anomaly	$\Delta x_1 = 30\text{m}$ $\Delta x_2 = 15\text{m}$			$\Delta x_1 = 30\text{m}$ $\Delta x_2 = 5\text{m}$			$\Delta x_1 = 15\text{m}$ $\Delta x_2 = 15\text{m}$			$\Delta x_1 = 15\text{m}$ $\Delta x_2 = 5\text{m}$		
	Width x Height x Length (m ³)	Depth (m)	ΔN (m)	ΔE (m)	Depth (m)	ΔN (m)	ΔE (m)	Depth (m)	ΔN (m)	ΔE (m)	Depth (m)	ΔN (m)
2x2x10	10	-13	-3	10	-13	-3	10	+2	-3	10	+2	+2
	10	-13	-3	20	-13	-3	10	+2	-3	20	+2	+2
	<i>10</i>	<i>+17</i>	<i>-3</i>	<i>10</i>	<i>-13</i>	<i>+2</i>	<i>10</i>	<i>+2</i>	<i>-3</i>	<i>20</i>	<i>+2</i>	<i>-3</i>
2x2x100	20	-13	-3	30	-13	-8	30	+2	-3	40	2	-8
	40	-13	-3	50	-13	-18	50	+2	-3	60	2	-13
	<i>30</i>	<i>-13</i>	<i>-18</i>	<i>50</i>	<i>-13</i>	<i>+2</i>	<i>40</i>	<i>+2</i>	<i>-3</i>	<i>60</i>	<i>+2</i>	<i>+7</i>
2x2x1000	50	-13	-18	70	-13	+87	70	+2	-18	80	+2	+7
	80	-13	-18	110	+17	-7	120	+2	-3	130	+17	-88
	<i>90</i>	<i>+17</i>	<i>-3</i>	<i>70</i>	<i>-13</i>	<i>+52</i>	<i>90</i>	<i>+2</i>	<i>-3</i>	<i>100</i>	<i>+2</i>	<i>+17</i>

NOTE:
Normal fonts indicate the detection results by the Γ_{33} matched filter with the psd model = $\sigma_w^2 \Delta x_1 \Delta x_2 [E^2/(\text{cyc/m})^2]$
Bold fonts indicate the detection results by the Γ_{33} matched filter with the empirical psd's for $1E^2/\text{Hz}$ white noise
Italic fonts indicate the detection results by the combination of the matched filters for all six gradient components using the psd models (assume no correlation between sensors).

Table 5.22: The anomaly detectability of the Γ_{33} matched filter using the psd model versus the empirical psd for $1E^2/\text{Hz}$ white noise (or $\sigma_w^2 = 1/3 E^2$ for the sampling frequency of $1/3 \text{ Hz}$) with respect to the cross-track spacing Δx_1 and the along-track data spacing Δx_2 . The matched filter for all six combined gradients are used in comparisons. The center of the anomaly at the azimuthal orientation of 90° is located at $N=18013\text{m}$ and $E=18018\text{m}$.

5.6 CASE V: Anomaly detectability versus orientation

In the previous sections, we only consider the matched filters fixed at the azimuthal orientations $\beta' = 90^\circ, 45^\circ, \text{ and } 0^\circ$ (in \mathbf{x} -triad frame in Figure B.1) to detect the 2m x 2m x 1000m anomaly having the same orientations, respectively, in noise environments. Since the component Γ_{33} due to the anomaly is invariant with respect to the orientations according to eq. (B.1), the Γ_{33} matched filter yields the best detection among other components. For this example, we investigate how sensitive the matched filters using a psd model are if they are rotated with respect to an anomaly in \mathbf{x} -triad frame. We test the gradient signals due to the 2m x 2m x 1000m anomaly added to $1E^2/\text{Hz}$ white noise (i.e., best-case detection scenario). The center of the anomaly is at N = 18000m and E=18000m. The anomaly has the azimuthal orientations, β' , of $5^\circ, 47^\circ, \text{ and } 86^\circ$ at 30m depth. Figure 5.24 shows the results of matched filtering in the noise. The matched filter using Psd model 4 is azimuthally oriented by 3° increments from 0° to 90° . For the anomaly at $\beta' = 5^\circ$, only the filters for the Γ_{33}, Γ_{22} and Γ_{23} components yield their the highest peaks close to the true orientation; the difference of the location at the highest peak minus the true location ΔN and ΔE approaching to zero in these components as shown in Figure 5.25.

For the anomaly at $\beta' = 47^\circ$, only the Γ_{33} and Γ_{12} matched filters give the highest peaks at 48° (Figure 5.24b) close to the true orientation where $\Delta N = 0\text{m}$ and $\Delta E = 0\text{m}$ (in Figure 5.26). For the anomaly at $\beta' = 86^\circ$, in Figure 5.24c, the plot shows that the anomaly is detectable using the Γ_{33}, Γ_{11} , and Γ_{13} components; see also Figure 5.27. Overall, the Γ_{33} matched filter gives the best anomaly detectability with respect to all orientations (i.e., $\beta' = 5^\circ, 47^\circ, \text{ and } 86^\circ$) of the anomaly.

As seen in Figure 5.28 the detection results are improved when the combination of all six combined gradients is used. The plot in Figure 5.28a shows that the highest peaks occur at 6° and are located at or near the true location of the anomaly at $\beta' = 5^\circ$, where ΔN and ΔE at 6° equal zero as can be seen in Figure 5.28d. The matched filters yield the best result for the multiple of 3° closest to the true orientation.

However, the results may deteriorate if the orientation increment increases. For example, if the increment is 6° , the highest peaks occurs at false locations offset by several kilometers with respect to the true location, see Figure 5.28a-f. For example, at the orientation of 12° in Figure 5.28a where the highest peak is 62.175, the differences are $\Delta N = -630\text{m}$ and $\Delta E = 4710\text{m}$ (Figure 5.28d). It can be explained that only when the filter is rotated close to the true orientation, the transfer function passes the frequencies of interest. For instance, figures 5.30b and 5.30e show the matched filter outputs in the frequency domain using the Γ_{33} matched filters at the false orientation (i.e., $\beta' = 0^\circ$) versus the correct orientation (i.e., $\beta' = 86^\circ$), respectively (figure 5.29 shows the characteristics of the gradient signals having $\beta' = 86^\circ$ and the observations in the space domain and the frequency domain). As such, the rotation can act as a band pass filter whose transfer function for different orientations scans through the frequencies of interest.

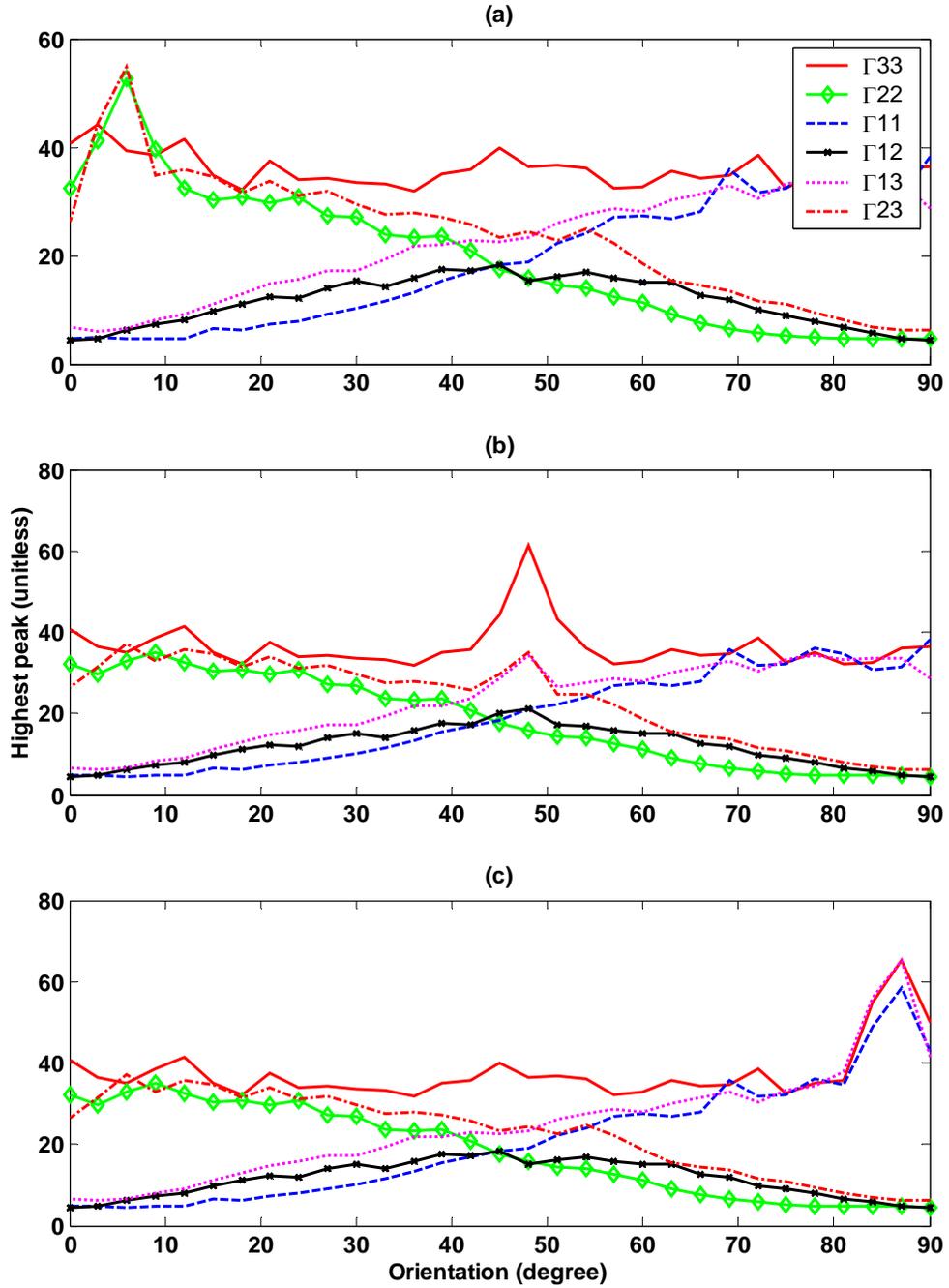


Figure 5.24: Anomaly detectability versus orientations in $1E^2/\text{Hz}$ white noise; the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly has the azimuthal orientations of (a) 05° (b) 47° and (c) 86° at 30m depth.

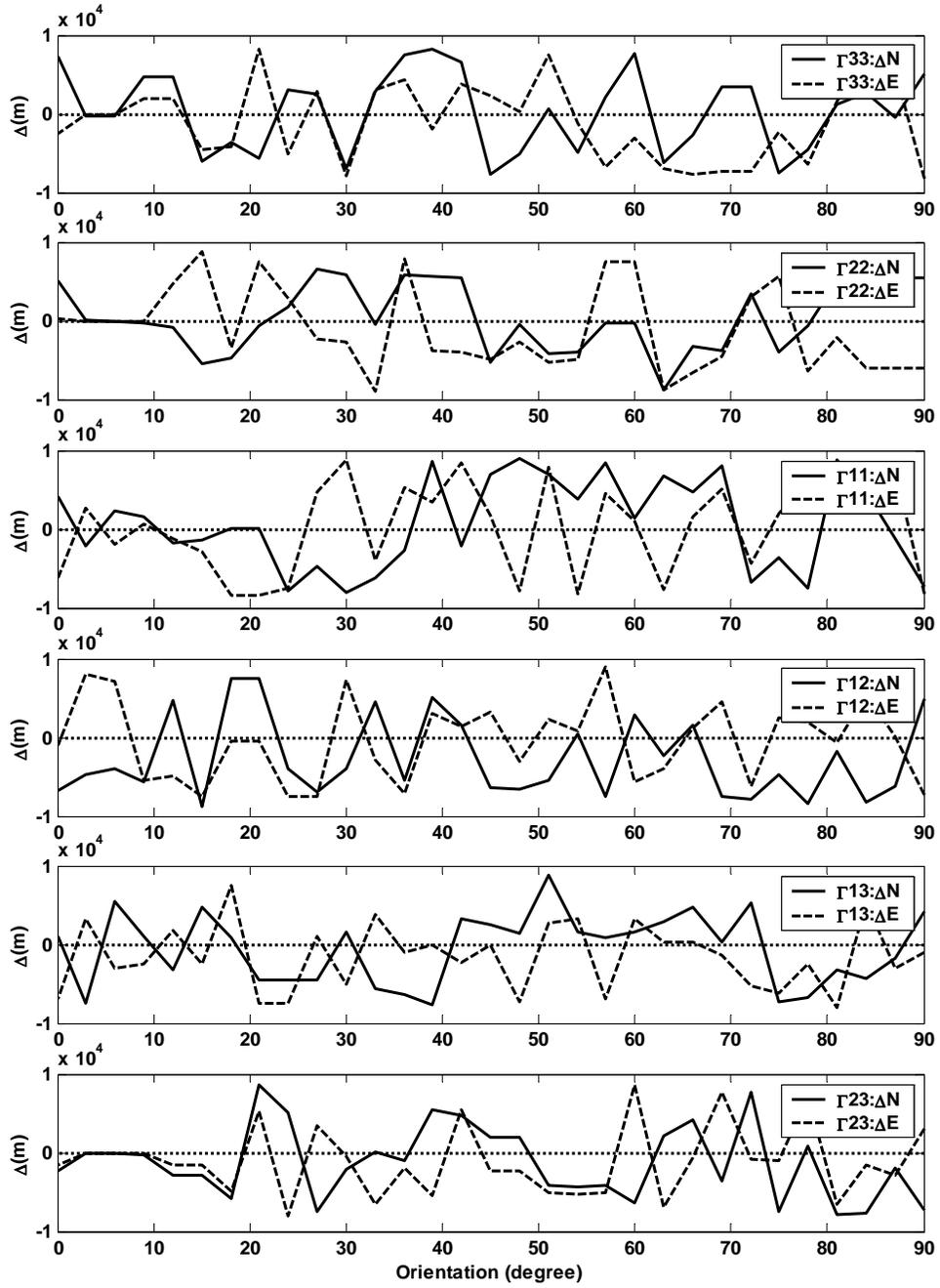


Figure 5.25: Anomaly detectability versus orientations in $1E^2/Hz$ white noise; the $2m \times 2m \times 1000m$ anomaly has the azimuthal orientation of 5° at $30m$ depth; ΔN and ΔE are the difference of the location of the highest peak minus the true location.

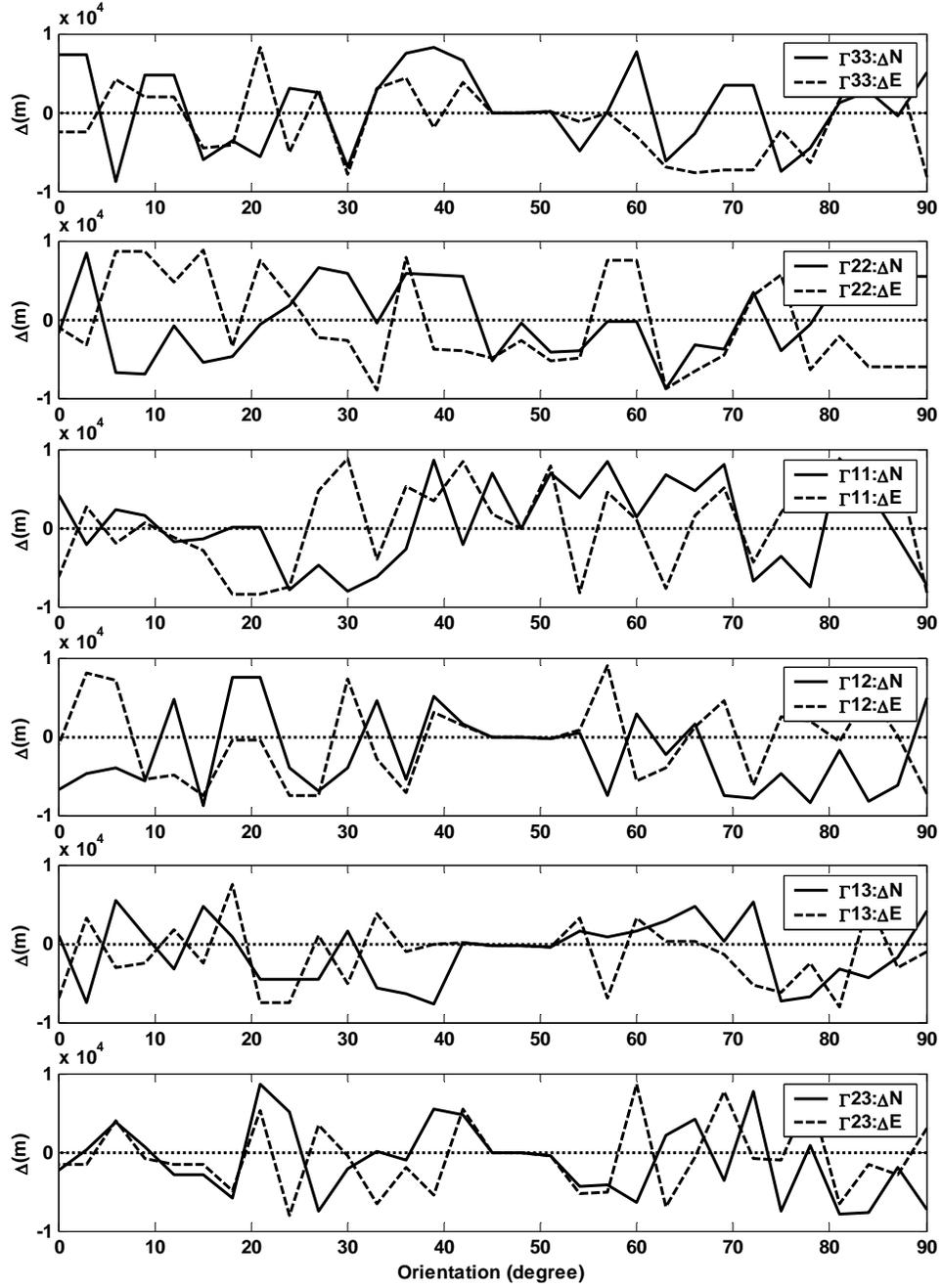


Figure 5.26: Anomaly detectability versus orientations in $1E^2/Hz$ white noise; the 2m x 2m x 1000m anomaly is has the azimuthal orientation of 47° at 30m depth; ΔN and ΔE are the difference of the location of the highest peak minus the true location.

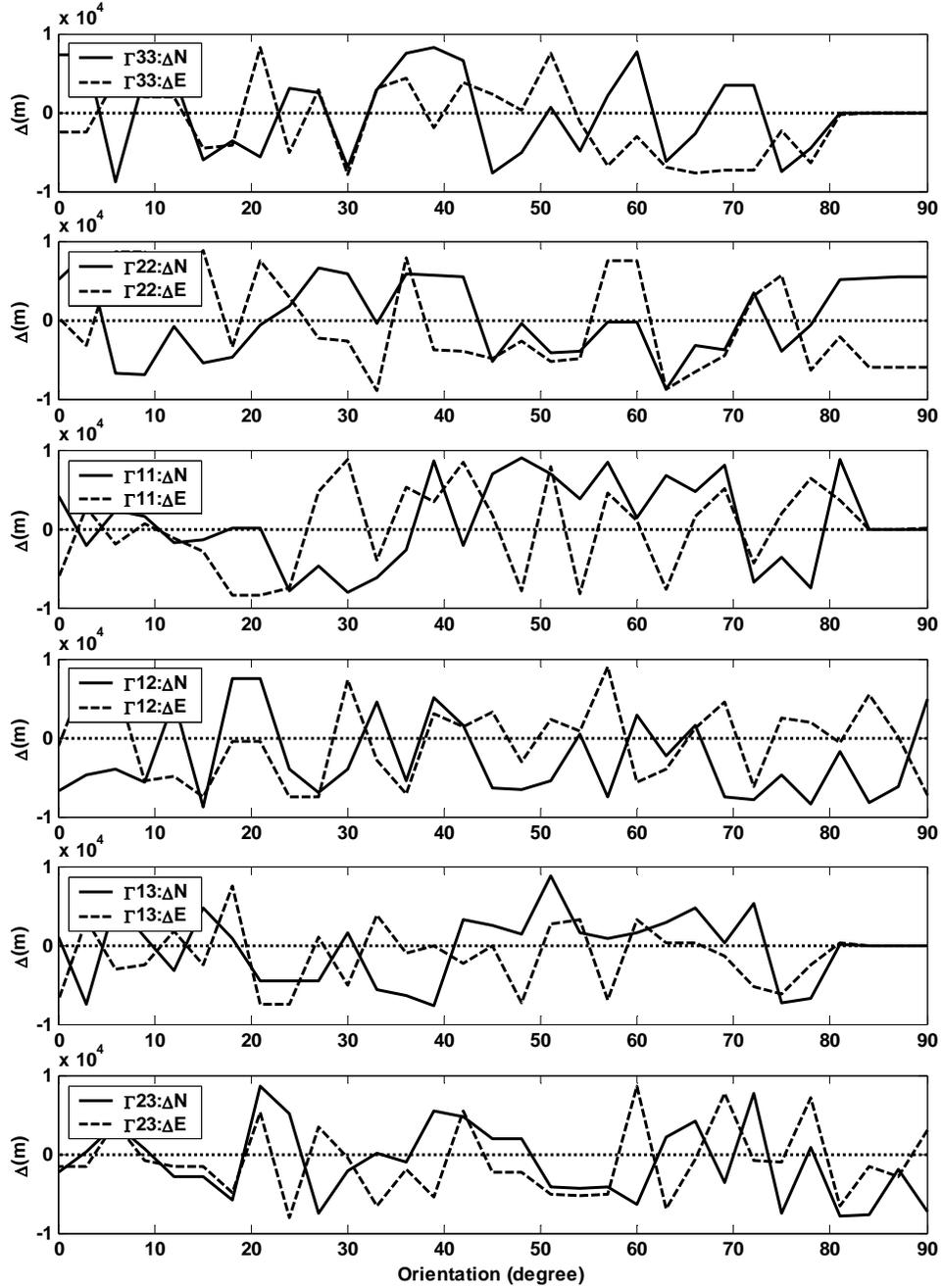


Figure 5.27: Anomaly detectability versus orientations in $1E^2/\text{Hz}$ white noise; the 2m x 2m x 1000m anomaly has the azimuthal orientation of 86° at 30m depth; ΔN and ΔE are the difference of the location of the highest peak minus the true location.

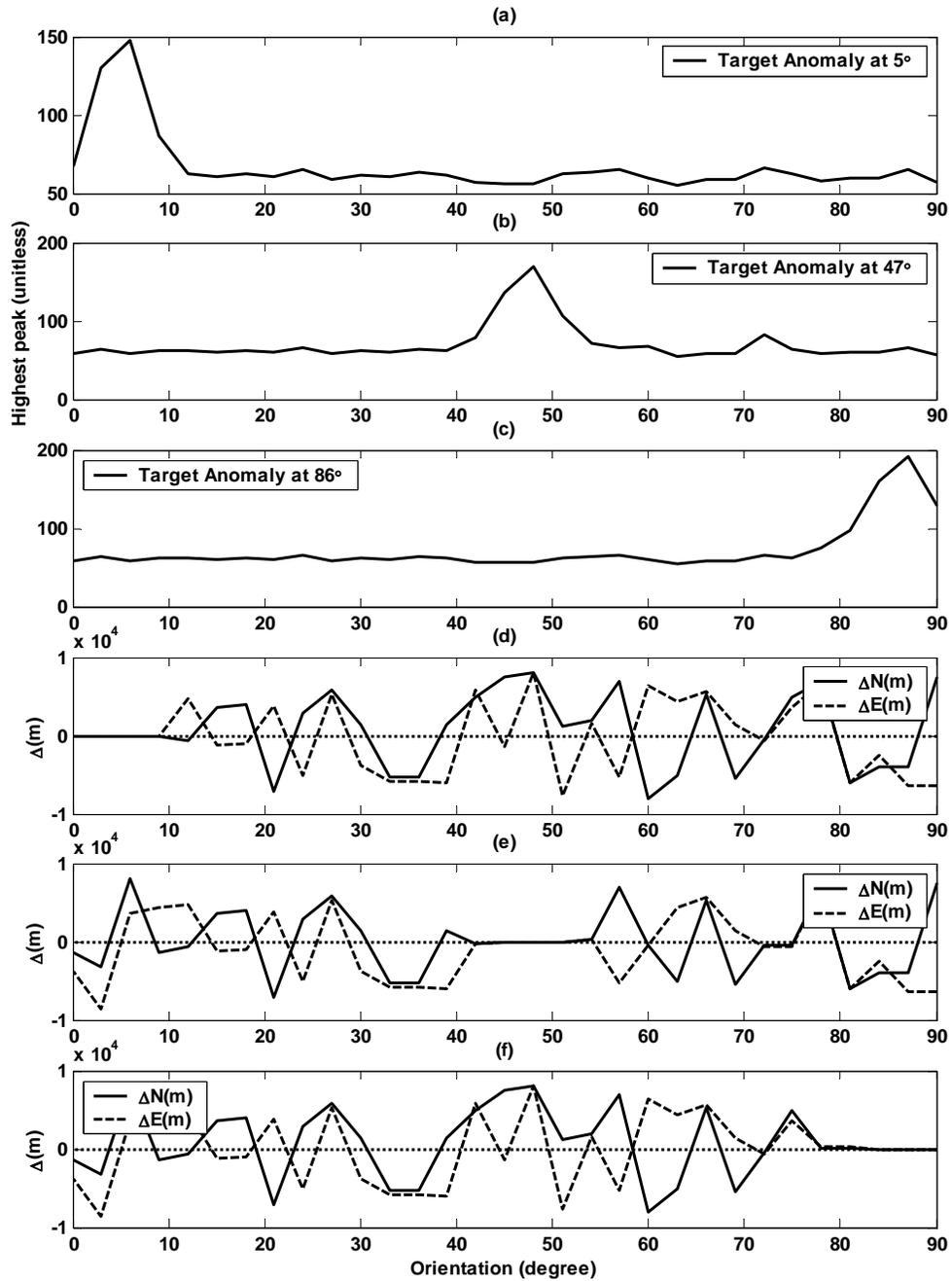


Figure 5.28: Anomaly detectability versus orientations in $1E^2/\text{Hz}$ white noise using all six combined gradients to detect the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly having the azimuthal orientations of (a) 5° (b) 47° and (c) 86° at 30m depth; (d), (e), and (f) show ΔN and ΔE with respect to (a), (b), and (c), respectively.

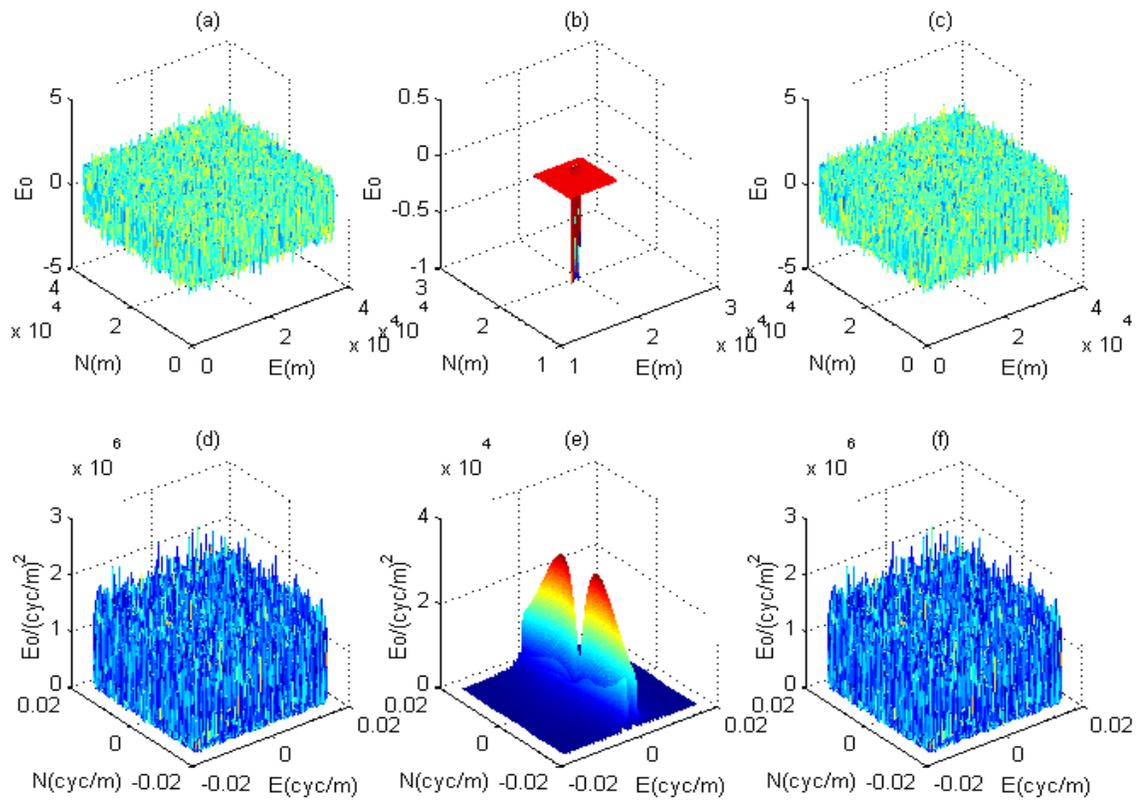


Figure 5.29: (a) $1E^2/\text{Hz}$ white noise; (b) The Γ_{33} gradient signal due to the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly having the azimuthal orientation of 86° ; (c) The observation: $1E^2/\text{Hz}$ white noise plus the Γ_{33} gradient signal; (d) The (magnitude) spectrum of (a); (e) The (magnitude) spectrum of (b); (f) The (magnitude) spectrum of (c).

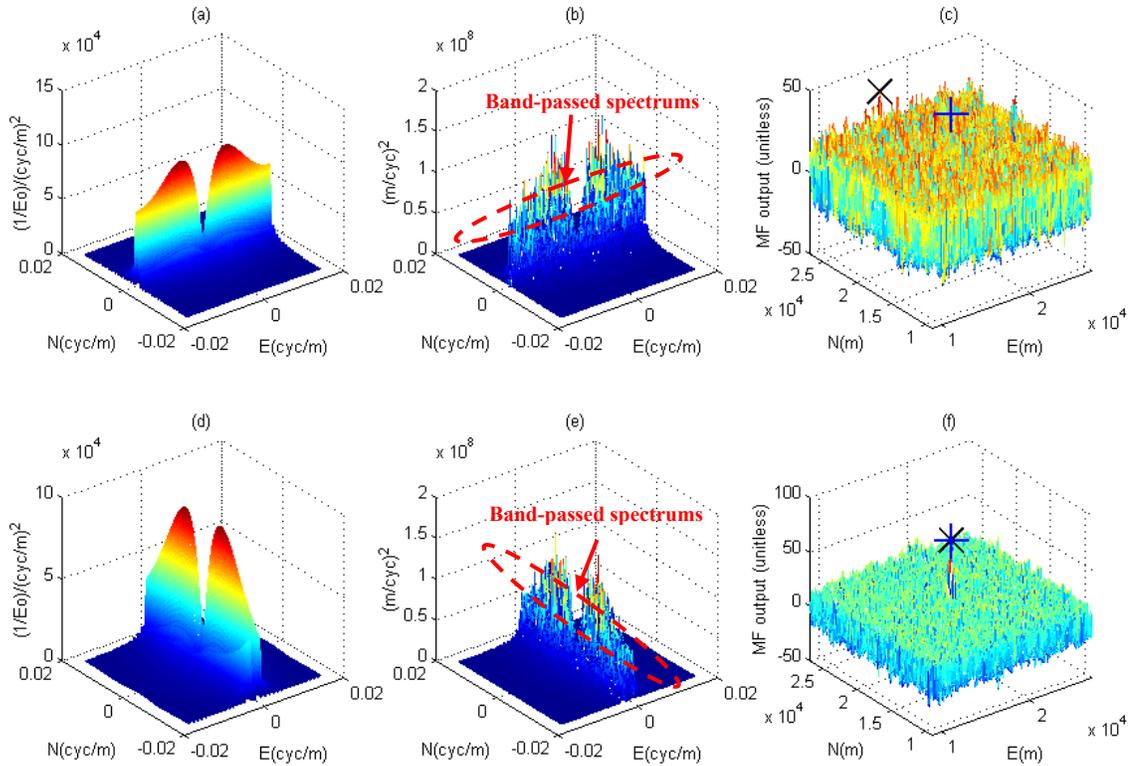


Figure 5.30: The upper row shows the case of using the false azimuthal orientation of 0° for the Γ_{33} matched filter to detect the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly having the azimuthal orientation, β' , of 86° ; (a) The transfer function of the matched filter (b) The (magnitude) spectrum of the matched filter output; (c) The highest peak output. The lower row shows the case of using the correct azimuthal orientation (i.e., $\beta' = 86^\circ$) for the Γ_{33} matched filter to detect the anomaly; (d) the transfer function of the matched filter; (e) The (magnitude) spectrum of the matched filter output; (f) The highest peak output. The symbol “x” indicates the location of the highest peak and “+” indicates the true location of the target anomaly.

5.7 Performance of matched filters

This section describes the detection performance of matched filters in various noise environments. We show the plots of the probability of false alarm versus depths for a given POM and different sizes and orientations of the target anomaly of interest. Based on the numerical results in CASE's I and II, we have chosen $POM = 0.001$ considered as a marginal probability level of the signal detection, since it effectively describes false locations by matched filtering, e.g., the detection results by using all six combined gradients in Table 5.10 (y_{\max} exceeds η' if POM is set too high).

Figures 5.31 – 5.33 show POF's versus depths for the matched filters using, respectively, the psd model versus the empirical psd for $1E^2/Hz$ white noise, with respect to 30m grid resolution and $\beta' = 0^\circ, 45^\circ, \text{ and } 90^\circ$. As described in CASE VI, the deepest detectable depths correspond to the detection performances portrayed by the plots in this section. The use of the empirical psd's indicates the increased detection performance, compared to the use of the psd model. For example, consider the 2m x 2m x 10m anomaly at and $\beta' = 90^\circ$ in $1E^2/Hz$ white noise, summarized in Table 5.16. The anomaly is detectable as deep as 20m by applying the Γ_{33} matched filter using the empirical psd, whereas only 10m if the psd model is used, see Table 5.13.

In Figure 5.31b, the detectable depth = 20m shows $POF = 0.264$. The failure of detection at 30m depth corresponds to $POF = 0.973$ which indicates a high chance of no signal at the location where the highest peak exists. It should be noted in Figure 5.32 that the orientation does not affect the performance of the Γ_{33} matched filter to detect, for example, the 2m x 2m x 100m anomaly. For all curves, POF rapidly increases at short ranges (10-30m) as depth increases, particularly for smaller sizes of the target anomaly.

Figure 5.34 provides the detection performance with respect to depths and sizes using the combination of all six gradients and the psd models for different types of noise backgrounds. The plots indicate the improvement of detection performance with the combination as compared to the case of the individual gradients shown in Figures 5.31(a) - 5.33(a). For example, the performance of the Γ_{33} matched filter using Psd model 4 for $1E^2/Hz$ white noise to detect the 2m x 2m x 100m anomaly at $\beta' = 0^\circ$ and 30m depth corresponds to $POF = 0.614$ (see Figure 5.32) while POF decreases to 0.097 using the combination (see Figure 5.34).

As can be seen in Figure 5.31, the detection performance for the target size of 2m x 2m x 10m (in $1E^2/Hz$ white noise), whose dimension is smaller than the spatial resolution of 30m, slightly increases when all six combined gradients are used (see the line “---” in Figure 5.35). Note that the higher the spatial resolution is, the more the SNR increases. Thus, the performance of matched filter also increases. For example, at a depth of 30m, POF decreases, roughly, 20% in association with the increase of the resolution from ($\Delta x_1 = 30m$ and $\Delta x_2 = 30m$; $POF = 0.982$) to ($\Delta x_1 = 15m$ and $\Delta x_2 = 5m$; $POF = 0.797$).

In summary, according to a number of numerical results from the previous cases, the detectability varies as the gradient signal due to the target anomaly attenuates with respect to the depth. The use of psd models slightly affects the matched filter detection. However, more accurate psd modeling is needed in order to enhance the capability of matched filter. Rather than doing so, the combination of all six combined gradients is an

alternative way to improve the detectability. The matched filters using empirical psd's show an advantage of signal detection because the characteristics of the psd's correspond to the noise signature. Furthermore, the location of the target anomaly affects the matched filter detectability.

In this study, although the performance of the matched filter can be improved by using the empirical psd's, the signal detectability still is limited if the signal signature strength (magnitude) is small, relative to strong noise backgrounds, e.g., the Γ_{12} gradient due to the 2m x 2m x 10m anomaly, shown in Figure 5.31. The spatial resolution is also a factor that affects the detectability which can be improved by considering higher resolution in an airborne gradiometric survey.

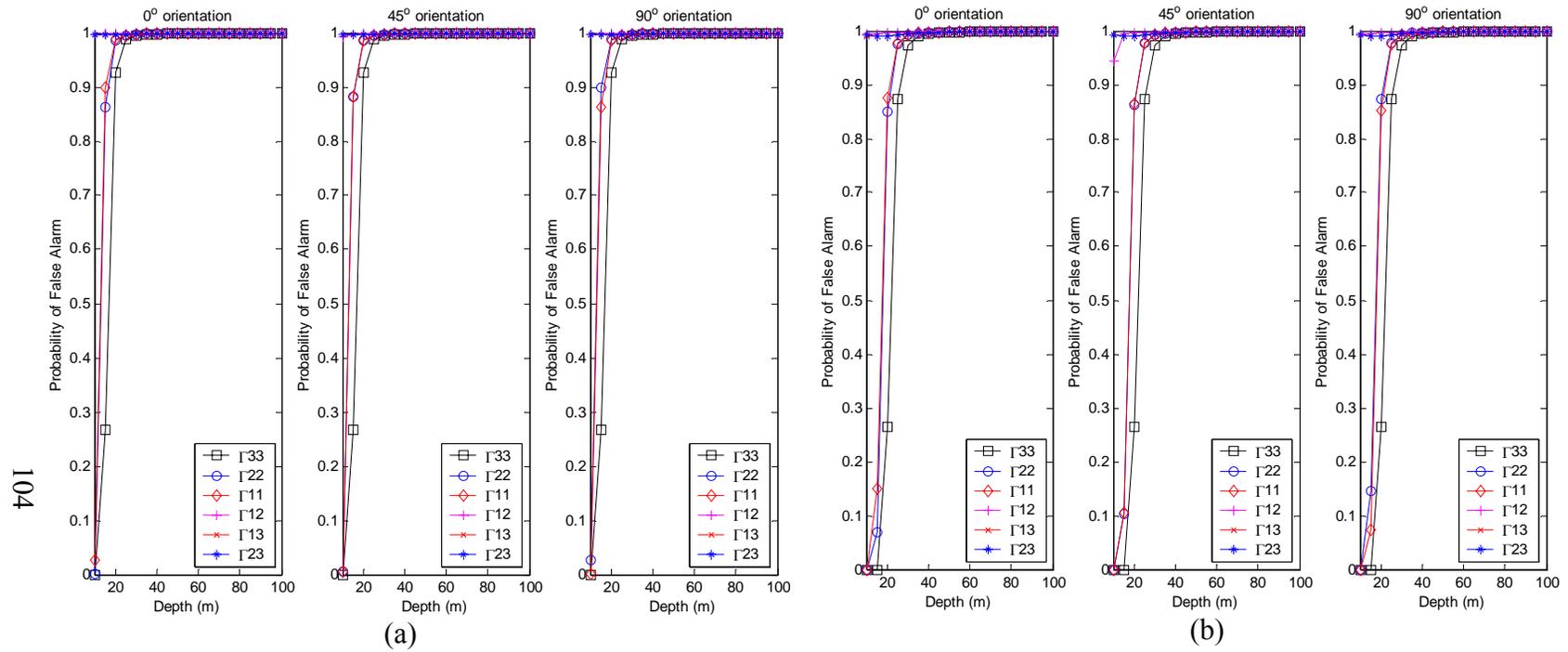


Figure 5.31: Performance of matched filters using (a) Psd model 1 and (b) the empirical psd for $1E^2/\text{Hz}$ white noise to detect the 2m x 2m x 10m anomaly.

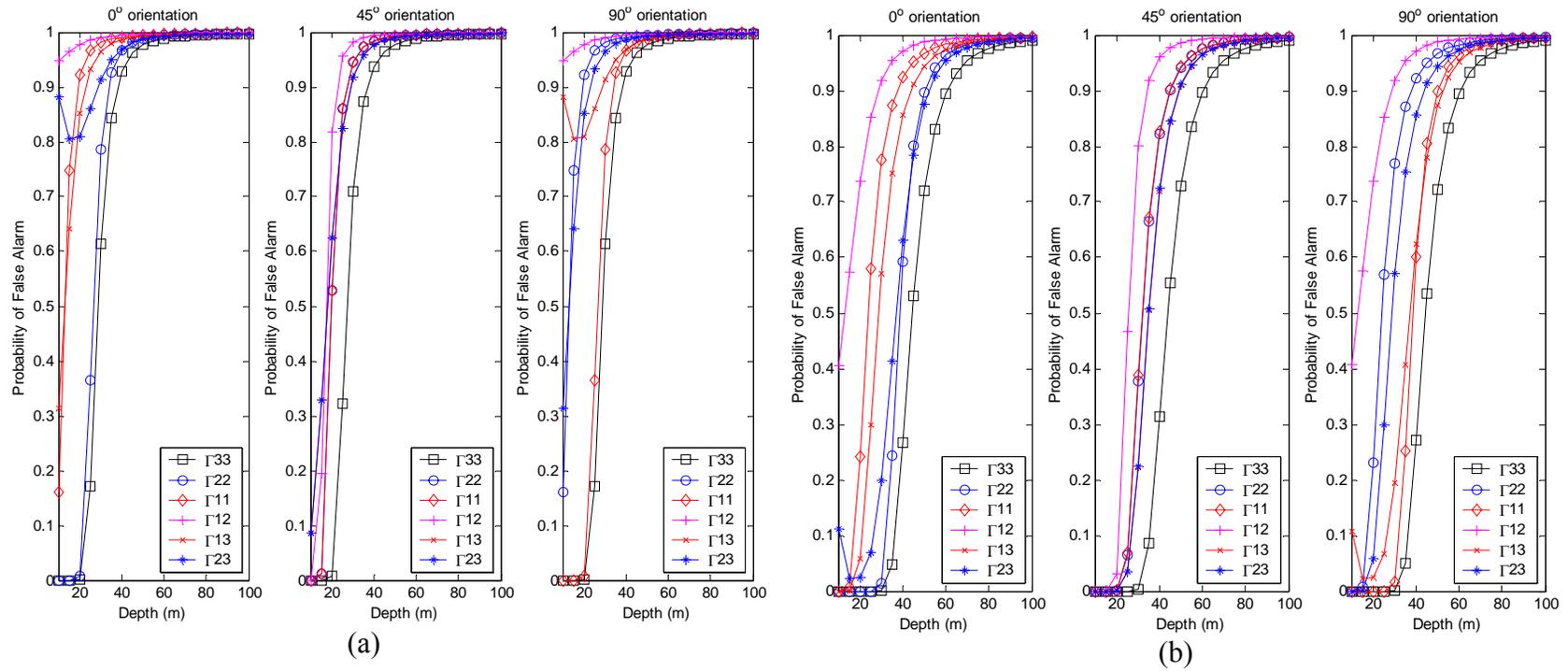


Figure 5.32: Performance of matched filters using (a) Psd model 1 and (b) the empirical psd for $1E^2/\text{Hz}$ white noise to detect the $2\text{m} \times 2\text{m} \times 100\text{m}$ anomaly

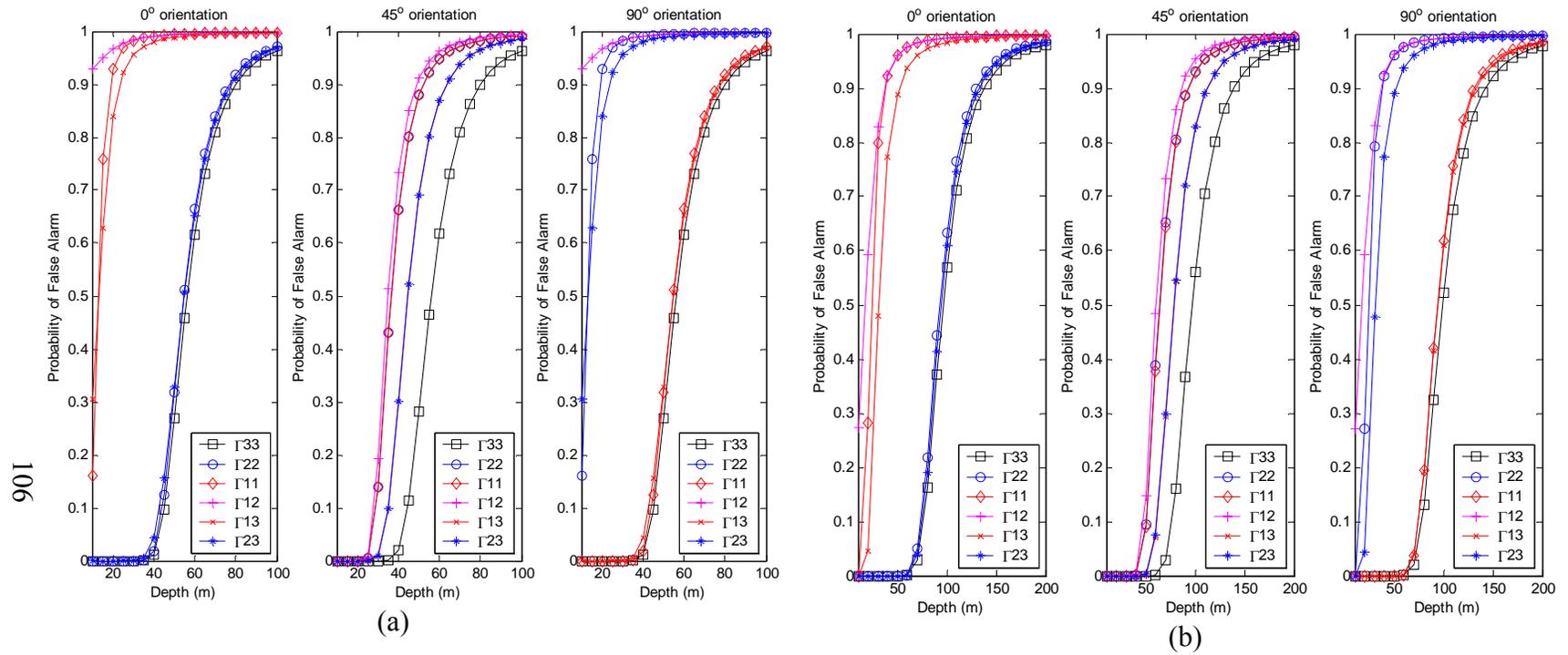


Figure 5.33: Performance of matched filters using (a) Psd model 1 and (b) the empirical psd for $1E^2/\text{Hz}$ white noise to detect the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly.

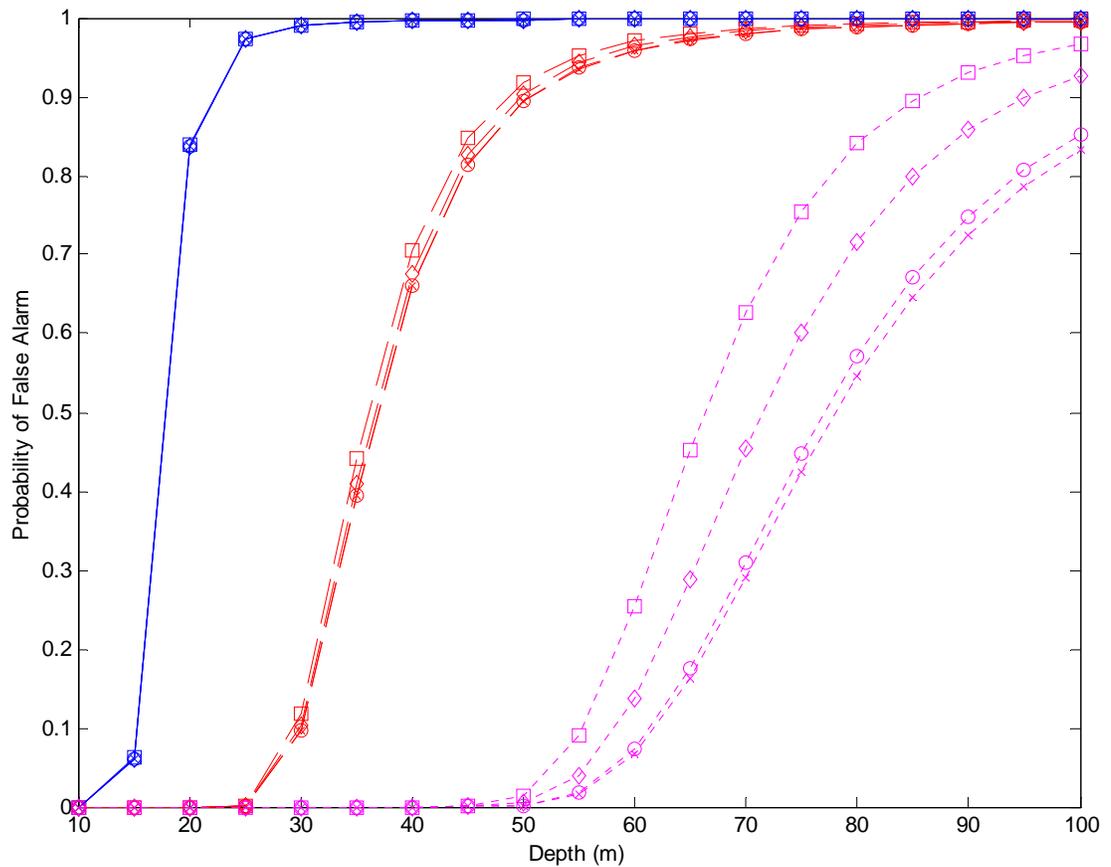


Figure 5.34: Performance of all six combined gradient matched filter; “□”= the full geological background plus $1E^2/Hz$ white noise; “◇” = the full geological background plus $1E^2/Hz$ white noise minus a geological model; “o” = the 5-km longer wavelength GGSS data plus $1E^2/Hz$ white noise; “x” = $1E^2/Hz$ white noise; solid line indicates the 2m x 2m x 10m anomaly; dashed line indicates the 2m x 2m x 100m anomaly; dotted line indicates the 2m x 2m x 1000m anomaly.

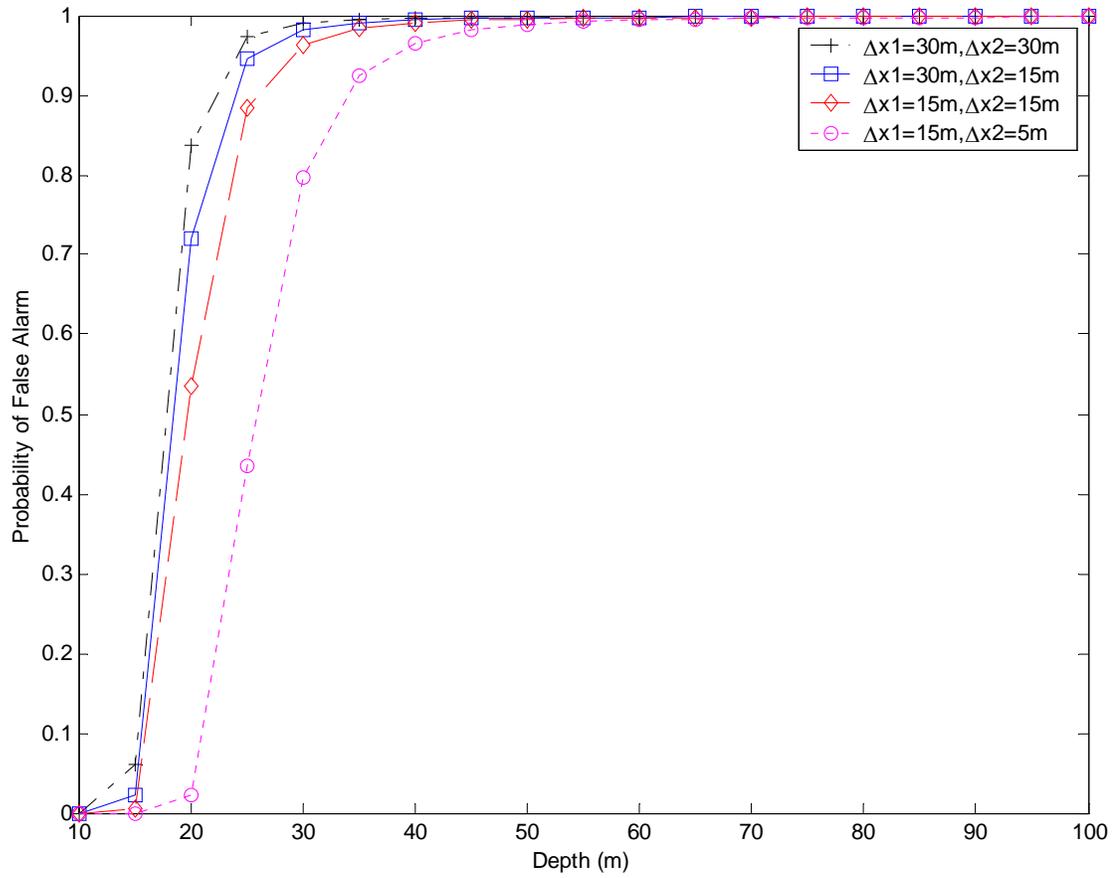


Figure 5.35: Performance of all six combined gradient matched filter for the 2m x 2m x 10m anomaly in $1E^2/\text{Hz}$ white noise with respect to spatial resolutions.

CHAPTER 6

THE THEORY OF PARAMETER ESTIMATION

6.1 Introduction

In previous chapters, we consider the gravitational gradients s_{ij} , with $(i, j) = (1, 2, 3)$, due to a mass anomaly, are exactly known except that the location of the anomaly is unknown. The matched filter is a function of every point of the observations, i.e., we wish to maximize the signal-to-noise ratio at the location, \bar{x} , of the anomaly. The most probable candidate(s) of the anomaly is chosen with a given probability of a miss at the point, where the highest filter output peak occurs. Throughout the filtering process, the matched filter requires a complete knowledge of the anomaly's signal so that the filter can accurately determine the location of the anomaly. In practice, we vaguely know about the anomalous source to be detected. Such a consequence greatly limits the performance of matched filters. The gradient signal s_{ij} , in turn, can be characterized by unknown parameters such as density contrast, 3D-location, shape, and orientation. The idea is to determine the parameters as accurately as possible and to use the corresponding estimates to reconstruct the matched filter. Therefore, when s_{ij} is not exactly known, the measured gradients becomes

$$v_{ij}(x) = s_{ij}(\boldsymbol{\theta}; x) + n_{ij}(x) \quad (6.1)$$

where $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m)$ is generally a $m \times 1$ vector of unknown parameters, including \bar{x} , to be determined.

In communications theory, the theory of parameter estimation mostly relies on certain assumptions of a-priori probability densities for the measurements (e.g., Gaussian type). They imply various types of parameter estimations, for instance, Bayes estimation, maximum a-posteriori estimation (equivalent to Bayes estimation after a quadratic cost function and a unimodal probability density function are applied, or known as minimum-mean-square-error estimation), and maximum likelihood estimation. More details can be found in Van Trees (1968), Whalen (1971), Poor (1983), and Kay (1993). For simplicity, we shall work on robust alternatives, which directly depend on a signal model, i.e., $s_{ij}(\boldsymbol{\theta})$. The estimation method of interest is based on the fundamental concept of least squares (Koch, 1999; Schaffrin, 1987, 1989, and 2001; Schaffrin and Bock, 1994):

$$[\mathbf{v} - \mathbf{s}_{ij}(\boldsymbol{\theta})]^T \boldsymbol{\Sigma}_n^{-1} [\mathbf{v} - \mathbf{s}_{ij}(\boldsymbol{\theta})] \rightarrow \min_{\boldsymbol{\theta}} \quad (6.2)$$

where $\boldsymbol{\Sigma}_n$ is a $n \times n$ positive definite covariance matrix of \mathbf{n} . The $n \times 1$ signal vector \mathbf{s}_{ij} is linearized and can be expressed as

$$\mathbf{s}_{ij}(\boldsymbol{\theta}) = \mathbf{s}_{ij}(\boldsymbol{\theta}_0 + \boldsymbol{\Theta}) = \mathbf{s}_{ij}(\boldsymbol{\theta}_0) + \mathbf{A}\boldsymbol{\Theta} \quad (6.3)$$

where \mathbf{A} is the $n \times m$ coefficient matrix associated with $\boldsymbol{\theta}$, $\boldsymbol{\theta}_0$ are a-priori (known) values of the parameters and $\boldsymbol{\Theta}$ are the “ m ” unknown (increment or perturbed) parameters; \mathbf{A} has the rank of m . In this study, we adopt this concept to the problem of spatial refinement of the anomaly location based on the random effects model (Schaffrin and Bock, 1994). The justification for using this approach is that the matched filter output has a statistical nature in the sense of random errors in the observations. Thus, the refinement of the anomaly location regarding to the perturbed parameters $\boldsymbol{\Theta}$ should be treated as “random effects” whose expectations are taken from a priori knowledge.

6.1 The random effects model and hypothesis testing

The approach treats $\boldsymbol{\theta}_0$ as random approximation of (non-random) parameters $\boldsymbol{\theta}$ and, therefore, the $\boldsymbol{\Theta}$ unknown parameters as “random effects (or random parameters)” whose first and second moments are given and considered as prior information. It should be kept in mind that the sum of $\boldsymbol{\theta}_0$ and $\boldsymbol{\Theta}$ is “non-random” so that $\boldsymbol{\theta}$ are “non-random” (see eq. (6.5)). The linearized model, which integrates stochastic prior information and actual observational information, is given by

$$\begin{aligned} \mathbf{y}_{ij} &= \mathbf{A}\boldsymbol{\Theta} + \mathbf{n}_{ij} \\ \boldsymbol{\Theta}_0 &= \boldsymbol{\Theta} + \mathbf{n}_{\Theta} \end{aligned} \quad \text{with} \quad \begin{bmatrix} \mathbf{n}_{ij} \\ \mathbf{n}_{\Theta} \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \sigma_0^2 \begin{bmatrix} \mathbf{C}_n & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\Theta} \end{bmatrix} \right) \quad (6.4)$$

where $\mathbf{y}_{ij} = \mathbf{v}_{ij} - \mathbf{s}_{ij}(\boldsymbol{\theta}_0)$. Our a-priori information is the vector of expected (bias) parameters, $\boldsymbol{\Theta}_0$, such that

$$\boldsymbol{\Theta}_0 = \varepsilon\{\boldsymbol{\theta} - \boldsymbol{\theta}_0\} = \boldsymbol{\theta} - \varepsilon\{\boldsymbol{\theta}_0\} \quad (6.5)$$

For instance, $\boldsymbol{\Theta}_0$ can be initially given as a zero vector and $\boldsymbol{\theta}_0$ can be the location \bar{x} of the anomaly obtained as a consequence of signal detection by matched filters. (It should be noted that the non-random $\mathbf{s}_{ij}(\boldsymbol{\theta})$, generated by a small anomalous mass, is entirely embedded in geological background noise; generally speaking, we may not know where the anomaly is in practice. Several matched filters may be needed for an appropriate choice of $\boldsymbol{\theta}_0$, e.g., good approximate values of \bar{x} . Thus, signal detection by matched filtering is useful, as a pre-processing stage contributing a priori information).

Generally, \mathbf{A} is a stochastic coefficient matrix due to the randomness of θ_0 , i.e., $\mathbf{A} = [\partial s_{ij}(\theta)/\partial \theta]_{\theta=\theta_0}$, and also $\varepsilon\{\mathbf{y}\} \neq \mathbf{A}\varepsilon\{\Theta\}$ due to the second and higher order terms of Θ (Schaffrin, 1985, pp. 285-300). The latter property leads to the requirement of higher-order terms in the Taylor series for the model. However, for simplicity in this study, we consider that $\varepsilon\{\mathbf{y}\} = \mathbf{A}\varepsilon\{\Theta\}$ throughout.

The positive-definite covariance (or dispersion) matrix, $\sigma_0^2 \mathbf{C}_n$ (or Σ_n), of zero-mean gradient noise (geophysical noise) vector \mathbf{n} is obtained from a reciprocal distance model of gravitational gradient (Jekeli, 2003), where \mathbf{C}_n is a cofactor matrix and σ_0^2 is the a-priori variance component. The vector of random errors \mathbf{n}_Θ has zero mean, with the positive-definite covariance matrix, $\Sigma_\Theta = \sigma_0^2 \mathbf{C}_\Theta$, given from signal detection processing. We also have $\Theta \sim (\Theta_0, \sigma_0^2 \mathbf{C}_\Theta)$. Since the prior information and the observation may be taken from different sources, we further assume the uncorrelated-ness of the respective random error vectors such that

$$\mathbf{C}\{\mathbf{n}, \mathbf{n}_\Theta\} = \varepsilon\{\mathbf{nn}_\Theta^T\} = [0] \quad (6.6)$$

The predicted parameters of Θ , which is called the Best inhomogeneously Linear Prediction (inhomBLIP), is given by

$$\begin{aligned} \tilde{\Theta} &= \Theta_0 + (\Sigma_\Theta^{-1} + \mathbf{A}^T \Sigma_n^{-1} \mathbf{A})^{-1} (\mathbf{A}^T \Sigma_n^{-1} \mathbf{y} - \mathbf{A}^T \Sigma_n^{-1} \mathbf{A} \Theta_0) \\ &= \Theta_0 + (\mathbf{C}_\Theta^{-1} + \mathbf{N})^{-1} \mathbf{N} (\mathbf{N}^{-1} \mathbf{C} - \Theta_0) \end{aligned} \quad (6.7)$$

together with the matrix of mean square prediction errors

$$\text{MSPE}\{\tilde{\Theta}\} = (\Sigma_\Theta^{-1} + \mathbf{A}^T \Sigma_n^{-1} \mathbf{A})^{-1} = \sigma_0^2 (\mathbf{C}_\Theta^{-1} + \mathbf{N})^{-1} \quad (6.8)$$

with $\mathbf{N} = \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A}$ and $\mathbf{C} = \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{y}$.

Let the predicted residuals be

$$\tilde{\mathbf{n}} = \mathbf{y} - \mathbf{A} \tilde{\Theta} \quad (6.9)$$

and

$$\tilde{\mathbf{n}}_\Theta = \Theta_0 - \tilde{\Theta} = -(\mathbf{C}_\Theta^{-1} + \mathbf{N})^{-1} \mathbf{N} (\mathbf{N}^{-1} \mathbf{C} - \Theta_0) \quad (6.10)$$

The estimate $\hat{\sigma}_0^2$ of σ_0^2 is

$$\hat{\sigma}_0^2 = \frac{\mathbf{\Omega}_R}{\varepsilon\{\mathbf{\Omega}_R / \sigma_0^2\}} \quad (6.11)$$

where

$$\begin{aligned}
\mathbf{\Omega}_R &= \tilde{\mathbf{n}}^T \mathbf{C}_n^{-1} \tilde{\mathbf{n}} + \tilde{\mathbf{n}}_\theta^T \mathbf{C}_\theta^{-1} \tilde{\mathbf{n}}_\theta \\
&= (\mathbf{y} - \mathbf{A}\tilde{\boldsymbol{\theta}})^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A}\tilde{\boldsymbol{\theta}}) + (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^T \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \\
&= \mathbf{y}^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A}\tilde{\boldsymbol{\theta}}) + \boldsymbol{\theta}_0^T \mathbf{C}_\theta^{-1} (\boldsymbol{\theta}_0 - \tilde{\boldsymbol{\theta}}) + \tilde{\boldsymbol{\theta}}^T (\mathbf{N}\tilde{\boldsymbol{\theta}} - \mathbf{C} + \mathbf{C}_\theta^{-1} \tilde{\boldsymbol{\theta}} - \mathbf{C}_\theta^{-1} \boldsymbol{\theta}_0) \\
&= \mathbf{y}^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A}\tilde{\boldsymbol{\theta}}) + \boldsymbol{\theta}_0^T \mathbf{C}_\theta^{-1} (\boldsymbol{\theta}_0 - \tilde{\boldsymbol{\theta}}) + \tilde{\boldsymbol{\theta}}^T (\mathbf{N}(\boldsymbol{\theta}_0 - \tilde{\mathbf{n}}_\theta) - \mathbf{C} - \mathbf{C}_\theta^{-1} \tilde{\mathbf{n}}_\theta) \\
&= \mathbf{y}^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A}\tilde{\boldsymbol{\theta}}) + \boldsymbol{\theta}_0^T \mathbf{C}_\theta^{-1} (\boldsymbol{\theta}_0 - \tilde{\boldsymbol{\theta}}) + \tilde{\boldsymbol{\theta}}^T (\mathbf{N}\boldsymbol{\theta}_0 - \mathbf{C} - (\mathbf{N} + \mathbf{C}_\theta^{-1}) \tilde{\mathbf{n}}_\theta) \\
&= \mathbf{y}^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A}\tilde{\boldsymbol{\theta}}) + \boldsymbol{\theta}_0^T \mathbf{C}_\theta^{-1} (\boldsymbol{\theta}_0 - \tilde{\boldsymbol{\theta}}) \\
&\quad + \tilde{\boldsymbol{\theta}}^T \underbrace{(\mathbf{N}\boldsymbol{\theta}_0 - \mathbf{C} + (\mathbf{N} + \mathbf{C}_\theta^{-1})(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1}(\mathbf{C} - \mathbf{N}\boldsymbol{\theta}_0))}_0 \\
&= \underbrace{\mathbf{y}^T \mathbf{C}_n^{-1} \mathbf{y} - \mathbf{C}^T \mathbf{N}^{-1} \mathbf{C} + \mathbf{C}^T \mathbf{N}^{-1} \mathbf{C} - \mathbf{C}^T \tilde{\boldsymbol{\theta}} + \boldsymbol{\theta}_0^T \mathbf{C}_\theta^{-1} (\boldsymbol{\theta}_0 - \tilde{\boldsymbol{\theta}})}_{\mathbf{\Omega}} \\
&= \mathbf{\Omega} + \underbrace{\mathbf{C}^T \mathbf{N}^{-1} \mathbf{C} - \mathbf{C}^T \tilde{\boldsymbol{\theta}} + \boldsymbol{\theta}_0^T \mathbf{C}_\theta^{-1} (\boldsymbol{\theta}_0 - \tilde{\boldsymbol{\theta}})}_{\mathbf{R}}
\end{aligned} \tag{6.12}$$

and $\varepsilon\{\mathbf{\Omega}_R/\sigma_0^2\} = n$, (see (6.20)). The quadratic form $\mathbf{\Omega}$ (unconstrained case) in (6.12) is given by

$$\begin{aligned}
\mathbf{\Omega} &= (\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\theta}})^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\theta}}) \\
&= \mathbf{y}^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A}\hat{\boldsymbol{\theta}}) - \hat{\boldsymbol{\theta}}^T \underbrace{(\mathbf{C} - \mathbf{N}\hat{\boldsymbol{\theta}})}_0 \\
&= \mathbf{y}^T \mathbf{C}_n^{-1} \mathbf{y} - \mathbf{C}^T \mathbf{N}^{-1} \mathbf{C}
\end{aligned} \tag{6.13}$$

with

$$\mathbf{N}\hat{\boldsymbol{\theta}} = \mathbf{C} \quad \text{and} \quad \hat{\boldsymbol{\theta}} = \mathbf{N}^{-1} \mathbf{C} \tag{6.14}$$

Using the definition of $\varepsilon\{\mathbf{a}^T \mathbf{B} \mathbf{a}\} = \text{tr}[\mathbf{B} \varepsilon\{\mathbf{a} \mathbf{a}^T\}]$, where the “tr” is the trace operator and “ \mathbf{a} ” and “ \mathbf{B} ” are an arbitrary vector and an arbitrary square matrix, the expectation of (6.13) is

$$\varepsilon\{\mathbf{\Omega}/\sigma_0^2\} = \sigma_0^{-2} \text{tr}[\mathbf{C}_n^{-1} \varepsilon\{\mathbf{y} \mathbf{y}^T\} - \mathbf{N}^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} \varepsilon\{\mathbf{y} \mathbf{y}^T\} \mathbf{C}_n^{-1} \mathbf{A}] = \text{tr}[\mathbf{I}_n - \mathbf{I}_m] = n - m \tag{6.15}$$

with $\varepsilon\{\mathbf{y} \mathbf{y}^T\} = \sigma_0^2 \mathbf{C}_n$. Consequently, from (6.10), the relative increase \mathbf{R} to the sum of squares $\mathbf{\Omega}$ of the residual \mathbf{n} after the inclusion of the stochastic constraint $\boldsymbol{\theta}_\theta = \boldsymbol{\theta} + \mathbf{n}_\theta$ in the linearized model (see also (6.4)) is obtained by

$$\begin{aligned}
\frac{\mathbf{R}}{\sigma_0^2} &= \sigma_0^{-2} (\boldsymbol{\Omega}_R - \boldsymbol{\Omega}) = \sigma_0^{-2} \mathbf{y}^T \mathbf{C}_n^{-1} (\mathbf{A}\mathbf{N}^{-1}\mathbf{C} - \mathbf{A}\tilde{\boldsymbol{\Theta}}) + \sigma_0^{-2} \boldsymbol{\Theta}_0^T \mathbf{C}_\theta^{-1} (\boldsymbol{\Theta}_0 - \tilde{\boldsymbol{\Theta}}) \\
&= \sigma_0^{-2} (\mathbf{A}\mathbf{N}^{-1}\mathbf{C} - \mathbf{A}\tilde{\boldsymbol{\Theta}})^T \mathbf{C}_n^{-1} (\mathbf{A}\mathbf{N}^{-1}\mathbf{C} - \mathbf{A}\tilde{\boldsymbol{\Theta}}) + \sigma_0^{-2} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)^T \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \\
&\quad + \underbrace{\sigma_0^{-2} (\mathbf{A}\tilde{\boldsymbol{\Theta}})^T \mathbf{C}_n^{-1} (\mathbf{A}\mathbf{N}^{-1}\mathbf{C} - \mathbf{A}\tilde{\boldsymbol{\Theta}}) + \tilde{\boldsymbol{\Theta}}^T \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)}_0 \\
&= \sigma_0^{-2} (\mathbf{N}^{-1}\mathbf{C} - \tilde{\boldsymbol{\Theta}})^T \mathbf{N} (\mathbf{N}^{-1}\mathbf{C} - \tilde{\boldsymbol{\Theta}}) + \sigma_0^{-2} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)^T \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \\
&= \sigma_0^{-2} \left((\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0) - (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{N} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0) \right)^T \mathbf{N} \left((\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0) - (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{N} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0) \right) \\
&\quad + \sigma_0^{-2} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0)^T \mathbf{N} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{N} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0) \\
&= \sigma_0^{-2} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)^T \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \\
&= \sigma_0^{-2} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0)^T \begin{bmatrix} \mathbf{N} - 2\mathbf{N}(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{N} \\ + \mathbf{N}(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{N} - \mathbf{C}_\theta^{-1}) (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{N} \\ + \mathbf{N}(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{N} \end{bmatrix} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0) \\
&= \sigma_0^{-2} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0)^T \left[\mathbf{N} - \mathbf{N}(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{N} \right] (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0) \\
&= \sigma_0^{-2} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0)^T (\mathbf{C}_\theta + \mathbf{N}^{-1})^{-1} (\mathbf{N}^{-1}\mathbf{C} - \boldsymbol{\Theta}_0)
\end{aligned} \tag{6.16}$$

or

$$\begin{aligned}
&= \sigma_0^{-2} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)^T (\mathbf{C}_\theta^{-1} + \mathbf{N}) \mathbf{N}^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1})^{-1} \mathbf{N}^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{N}) (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \\
&= \sigma_0^{-2} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)^T \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)
\end{aligned}$$

It can be shown that the term in “ $\underbrace{\hspace{2cm}}$ ” of (6.16) is zero:

$$\begin{aligned}
&(\mathbf{A}\tilde{\boldsymbol{\Theta}})^T \mathbf{C}_n^{-1} (\mathbf{A}\mathbf{N}^{-1}\mathbf{C} - \mathbf{A}\tilde{\boldsymbol{\Theta}}) + \tilde{\boldsymbol{\Theta}}^T \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) = \\
&= (\mathbf{A}\tilde{\boldsymbol{\Theta}})^T \mathbf{C}_n^{-1} (\mathbf{A}\mathbf{N}^{-1}\mathbf{C} - \mathbf{A}\boldsymbol{\Theta}_0 - \mathbf{A}(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} (\mathbf{C} - \mathbf{N}\boldsymbol{\Theta}_0)) + \tilde{\boldsymbol{\Theta}}^T \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \\
&= \tilde{\boldsymbol{\Theta}}^T \left(\mathbf{N}\mathbf{N}^{-1}\mathbf{C} - \mathbf{N}\boldsymbol{\Theta}_0 - \mathbf{N}(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} (\mathbf{C} - \mathbf{N}\boldsymbol{\Theta}_0) \right) + \tilde{\boldsymbol{\Theta}}^T \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \\
&= \tilde{\boldsymbol{\Theta}}^T \left(\mathbf{I}_m - \mathbf{N}(\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \right) (\mathbf{C} - \mathbf{N}\boldsymbol{\Theta}_0) + \tilde{\boldsymbol{\Theta}}^T \mathbf{C}_\theta^{-1} \underbrace{(\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)}_{-\tilde{\mathbf{n}}_\theta}
\end{aligned}$$

$$\begin{aligned}
&= \tilde{\boldsymbol{\Theta}}^T \left(\mathbf{A}^T \mathbf{C}_n^{-1} - \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A} (\mathbf{C}_\theta^{-1} + \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} \right) (\mathbf{y} - \mathbf{A} \boldsymbol{\Theta}_0) \\
&- \tilde{\boldsymbol{\Theta}}^T \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{A} \boldsymbol{\Theta}_0) \\
&= \tilde{\boldsymbol{\Theta}}^T \left(\begin{array}{c} \mathbf{A}^T \mathbf{C}_n^{-1} - \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A} (\mathbf{C}_\theta^{-1} + \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} \\ - \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} \end{array} \right) (\mathbf{y} - \mathbf{A} \boldsymbol{\Theta}_0) \\
&= \tilde{\boldsymbol{\Theta}}^T \underbrace{\left(\mathbf{A}^T \mathbf{C}_n^{-1} - (\mathbf{C}_\theta^{-1} + \mathbf{A}^T \mathbf{C}_n^{-1} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} \right)}_0 (\mathbf{y} - \mathbf{A} \boldsymbol{\Theta}_0) \\
&= 0
\end{aligned} \tag{6.17}$$

Clearly,

$$\mathbf{R} \geq 0 \tag{6.18}$$

The expectation of \mathbf{R}/σ_0^2 is

$$\begin{aligned}
\varepsilon \left\{ \frac{\mathbf{R}}{\sigma_0^2} \right\} &= \sigma_0^{-2} \varepsilon \left\{ (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)^T \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \right\} \\
&= \sigma_0^{-2} \text{tr} \left[\mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} \varepsilon \{ \tilde{\mathbf{n}}_\theta \tilde{\mathbf{n}}_\theta^T \} \right] \\
&= \sigma_0^{-2} \text{tr} \left[\mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} \varepsilon \{ (\mathbf{y} - \mathbf{A} \boldsymbol{\Theta}_0) (\mathbf{y} - \mathbf{A} \boldsymbol{\Theta}_0)^T \} \mathbf{C}_n^{-1} \mathbf{A} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \right] \\
&= \text{tr} \left[\mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} \left((\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} (\mathbf{C}_n + \mathbf{A} \mathbf{C}_\theta \mathbf{A}^T) \mathbf{C}_n^{-1} \mathbf{A} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \right) \right] \\
&= \text{tr} \left[\mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} \left((\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} (\mathbf{N} + \mathbf{N} \mathbf{C}_\theta \mathbf{N}) (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \right) \right] \\
&= \text{tr} \left[\mathbf{C}_\theta^{-1} (\mathbf{C}_\theta + \mathbf{N}^{-1}) \mathbf{C}_\theta^{-1} \left(\mathbf{C}_\theta (\mathbf{I}_m + \mathbf{N} \mathbf{C}_\theta)^{-1} (\mathbf{I}_m + \mathbf{N} \mathbf{C}_\theta) \mathbf{N} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \right) \right] \\
&= \text{tr} \left[(\mathbf{C}_\theta + \mathbf{N}^{-1}) (\mathbf{C}_\theta^{-1} + \mathbf{N} - \mathbf{C}_\theta^{-1}) (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{C}_\theta^{-1} \right] \\
&= \text{tr} \left[(\mathbf{C}_\theta + \mathbf{N}^{-1}) (\mathbf{C}_\theta^{-1} - \mathbf{C}_\theta^{-1} (\mathbf{C}_\theta^{-1} + \mathbf{N})^{-1} \mathbf{C}_\theta^{-1}) \right] \\
&= \text{tr} \left[(\mathbf{C}_\theta + \mathbf{N}^{-1}) (\mathbf{C}_\theta + \mathbf{N}^{-1})^{-1} \right] \\
&= \text{tr} [\mathbf{I}_m] \\
&= m
\end{aligned} \tag{6.19}$$

Thus, with (6.15) and (6.19),

$$\varepsilon \{ \boldsymbol{\Omega}_R / \sigma_0^2 \} = \varepsilon \{ (\boldsymbol{\Omega} + \mathbf{R}) / \sigma_0^2 \} = \mathbf{n} \tag{6.20}$$

For the situation that a gradient signal generated by an anomaly to be detected is in strong geological noise, we may consider whether or not the prior information, which is the output of detection by matched filtering, is consistent with the actual data (i.e., \mathbf{y}).

Thus, we proceed with hypothesis testing:

$$H_0 : \varepsilon\{\Theta\} - \Theta_0 = \varepsilon\{\Theta - \Theta_0\} = \mathbf{0} \quad (6.21)$$

versus

$$H_1 : \varepsilon\{\Theta\} - \Theta_0 = \varepsilon\{\Theta - \Theta_0\} = \boldsymbol{w}_0 \neq \mathbf{0} \quad (6.22)$$

Note that the parameters Θ and σ_0^2 are unknown and the expectation of $\tilde{\Theta}$ in (6.7) is

$$\begin{aligned} \varepsilon\{\tilde{\Theta}\} &= \varepsilon\left\{\Theta_0 + (\mathbf{C}_\Theta^{-1} + \mathbf{N})^{-1} \mathbf{N}(\mathbf{N}^{-1} \mathbf{C} - \Theta_0)\right\} \\ &= \Theta_0 + (\mathbf{C}_\Theta^{-1} + \mathbf{N})^{-1} \mathbf{N}(\mathbf{N}^{-1} \mathbf{A}^T \mathbf{C}_n^{-1} \varepsilon\{\mathbf{y} = \mathbf{A}\Theta + \mathbf{n}\} - \Theta_0) \\ &= \Theta_0 + (\mathbf{C}_\Theta^{-1} + \mathbf{N})^{-1} \mathbf{N} \underbrace{\varepsilon\{\Theta\} - \Theta_0}_0 \\ &= \varepsilon\{\Theta\} \end{aligned} \quad (6.23)$$

Equivalently, we may test (based on Schaffrin (2001))

$$H_0 : \varepsilon\{\tilde{\Theta}\} - \Theta_0 = \varepsilon\{\tilde{\Theta} - \Theta_0\} = \mathbf{0} \quad (6.24)$$

versus

$$H_1 : \varepsilon\{\tilde{\Theta}\} - \Theta_0 = \varepsilon\{\tilde{\Theta} - \Theta_0\} = \boldsymbol{w}_0 \neq \mathbf{0} \quad (6.25)$$

with

$$\begin{aligned} D\{\tilde{\Theta} - \Theta_0\} &= \varepsilon\left\{(\tilde{\Theta} - \Theta_0)(\tilde{\Theta} - \Theta_0)^T\right\} \\ &= \Sigma_\Theta \left(\Sigma_\Theta + (\mathbf{A}^T \Sigma_n^{-1} \mathbf{A})^{-1}\right)^{-1} \Sigma_\Theta \\ &= \sigma_0^2 \mathbf{C}_\Theta (\mathbf{C}_\Theta + \mathbf{N}^{-1})^{-1} \mathbf{C}_\Theta \end{aligned} \quad (6.26)$$

The test statistic relating to this set of hypotheses can be expressed as (Koch, 1999, p.271-280):

$$\begin{aligned} T &= \frac{(\mathbf{R}/\sigma_0^2)/\varepsilon\{\mathbf{R}/\sigma_0^2\}}{(\mathbf{\Omega}/\sigma_0^2)/\varepsilon\{\mathbf{\Omega}/\sigma_0^2\}} \\ &= \frac{(\tilde{\Theta} - \Theta_0)^T [\sigma_0^{-2} \mathbf{C}_\Theta^{-1} (\mathbf{C}_\Theta + \mathbf{N}^{-1}) \mathbf{C}_\Theta^{-1}] (\tilde{\Theta} - \Theta_0) / m}{\sigma_0^{-2} \mathbf{\Omega} / (n - m)} \\ &= \frac{1}{m \bar{\sigma}_0^2} (\tilde{\Theta} - \Theta_0)^T \mathbf{C}_\Theta^{-1} (\mathbf{C}_\Theta + \mathbf{N}^{-1}) \mathbf{C}_\Theta^{-1} (\tilde{\Theta} - \Theta_0) \end{aligned} \quad (6.27)$$

where $\bar{\sigma}_0^2 = \mathbf{\Omega} / (n - m)$ is the estimate of σ_0^2 without the stochastic constraint of (6.4). We immediately see the test contains the inverted term of (6.26). That is, T is significantly greater than zero if H_1 is true. The distribution of the test statistic follows a central F-distribution

$$T \sim F(m, n - m) \tag{6.28}$$

if the null hypothesis H_0 is true--it means the prior information is consistent with the actual data. Otherwise,

$$T \sim F(m, n - m; \lambda) \tag{6.29}$$

with the non-centrality parameter defined as (see also Koch, 1999, p. 126-131):

$$\lambda = \sigma_0^{-2} w_0^T C_\theta^{-1} (C_\theta + N^{-1}) C_\theta^{-1} w_0 \tag{6.30}$$

if the alternative hypothesis with specified values of w_0 and σ_0^2 is correct—it means the prior information is not consistent with the actual data.

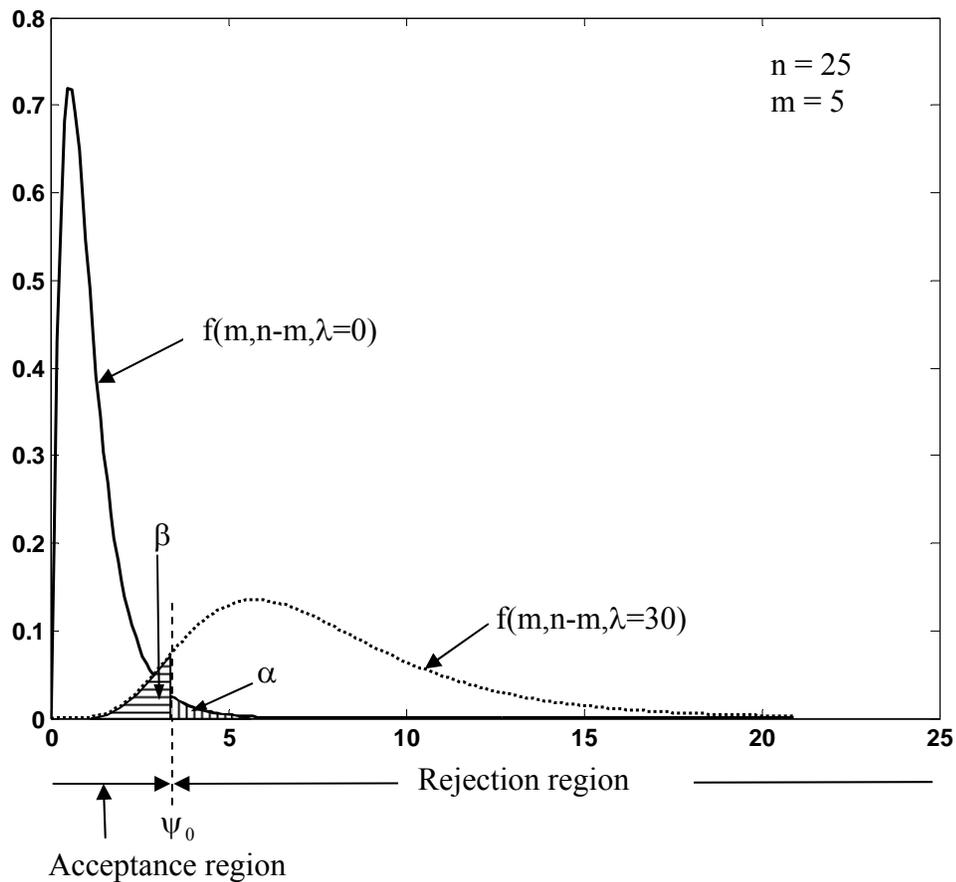


Figure 6.1: Acceptance and rejection region.

As shown in Figure 6.1, the test statistic basically follows the traditional F-test (centrality versus non-centrality). With a significance level α defined as the probability of a Type I error (i.e., we accept H_1 but H_0 is true (Jekeli, 2006)), we accept H_0 (i.e., T is below the threshold ψ_0) rather than H_1 . Consequently, the probability of a Type II error, β , (i.e., we accept H_0 but H_1 is true) can be determined based on ψ_0 and the non-central parameter, λ , of (6.30). The error probabilities are defined as

$$\alpha = 1 - \int_0^{\psi_0} f(m, n - m; \lambda = 0) d\psi \quad (6.31)$$

and

$$\beta = \int_0^{\psi_0} f(m, n - m; \lambda) d\psi \quad (6.32)$$

where f is the probability density function of F-distribution and ψ_0 is computed based on (6.31) with a chosen value of α .

6.3 Implementation

In this section, we apply the random effects model for the case of single and multiple data tracks with single and multiple sensors. A main concern of using the approach is the problem of matrix inversion, especially, the inverse of the large covariance matrix. In order to deal with the problem, we introduce a transformation matrix, which is orthogonal, and apply it to the first observation equation in (6.4) and its covariance matrix. As a consequence, the corresponding covariance matrix becomes diagonal and can be inverted simply—only diagonal elements or blocks are inverted. The following is a brief discussion of the derivations and implementation for the prediction of parameters, $\tilde{\Theta}$, together with their covariance matrix. More details of derivations can be found in Appendix C.

6.3.1 Single data track and single sensor (SS)

The $N \times 1$ observation vector \mathbf{y} can be transformed by using an orthogonal matrix, \mathbf{F} , such that $\mathbf{F}^H \mathbf{F} = \mathbf{I}$ and $\mathbf{F}^H = \mathbf{F}^{-1}$ where the ‘‘H’’ is complex conjugate transpose. The orthogonal matrix containing the N orthonormal basis vectors $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}$ is defined as (Kay, 1998, p.38-35):

$$\mathbf{F} = [\mathbf{v}_0 \quad \mathbf{v}_1 \quad \dots \quad \mathbf{v}_{N-1}] \quad (6.33)$$

with

$$\mathbf{v}_p = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 \\ e^{i2\pi\Delta x f_p} \\ \vdots \\ e^{i2\pi\Delta x (N-1)f_p} \end{bmatrix} \quad (6.34)$$

and

$$\mathbf{v}_{N-p} = \mathbf{v}_p^* \quad (6.35)$$

where $f_p = \frac{p}{N\Delta x}$ for $p = 0, 1, 2, \dots, N-1$, the asterisk denotes complex conjugate, and Δx is a sampling interval. It can be shown that the normalized exponential vectors are orthogonal as follows

$$\begin{aligned} \mathbf{v}_p^H \mathbf{v}_q &= \frac{1}{N} \begin{bmatrix} 1 & e^{-i2\pi\Delta x f_p} & \dots & e^{-i2\pi(N-1)\Delta x f_p} \end{bmatrix} \begin{bmatrix} 1 \\ e^{i2\pi\Delta x f_q} \\ \vdots \\ e^{i2\pi(N-1)\Delta x f_q} \end{bmatrix} \\ &= \frac{1}{N} \left(1 + e^{-i2\pi\Delta x(f_p - f_q)} + \dots + e^{-i2\pi(N-1)\Delta x(f_p - f_q)} \right) \\ &= \frac{1}{N} \left(\sum_{k=0}^{N-1} e^{-i2\pi k \Delta x(f_p - f_q)} \right) \\ &= \begin{cases} 1 & \text{if } f_p = f_q \\ 0 & \text{if } f_p \neq f_q \end{cases} \end{aligned} \quad (6.36)$$

By multiplying the orthogonal matrix \mathbf{F}^H on both sides of \mathbf{y} in (6.4), we have

$$\tilde{\mathbf{y}}_{SS} = \tilde{\mathbf{A}}_{SS} \boldsymbol{\Theta} + \tilde{\mathbf{n}}_{SS} \quad (6.37)$$

where $\tilde{\mathbf{y}}_{SS} = \mathbf{F}^H \mathbf{y}$, $\tilde{\mathbf{A}}_{SS} = \mathbf{F}^H \mathbf{A}$, and $\tilde{\mathbf{n}}_{SS} = \mathbf{F}^H \mathbf{n}$ and the subscript ‘‘SS’’ stands for ‘‘Single data track and Single sensor’’. For a convenience, let $\boldsymbol{\Sigma}_{SS} = \boldsymbol{\Sigma}_n$, the corresponding psd becomes diagonal and is given by

$$\tilde{\boldsymbol{\Sigma}}_{SS} = \mathbf{F}^H \boldsymbol{\Sigma}_{SS} \mathbf{F} = \text{diag}(\tilde{\boldsymbol{\Sigma}}_{SS}(f_0), \tilde{\boldsymbol{\Sigma}}_{SS}(f_1), \dots, \tilde{\boldsymbol{\Sigma}}_{SS}(f_1), \dots, \tilde{\boldsymbol{\Sigma}}_{SS}(f_{N-1})) \text{ with} \quad (6.38)$$

$$\tilde{\boldsymbol{\Sigma}}_{SS}(f_i) = \sum_{k=0}^{N-1} r(k) e^{-i2\pi k \Delta x f_i} \quad (6.39)$$

According to the definition of \mathbf{F} , the predicted parameters of $\boldsymbol{\Theta}$ in (6.7) can be expressed as

$$\begin{aligned}
\tilde{\boldsymbol{\Theta}} &= \boldsymbol{\Theta}_0 + \left(\boldsymbol{\Sigma}_\Theta + \mathbf{A}^T \boldsymbol{\Sigma}_{SS}^{-1} \mathbf{A} \right)^{-1} \left(\mathbf{A}^T \boldsymbol{\Sigma}_{SS}^{-1} \mathbf{y} - \mathbf{A}^T \boldsymbol{\Sigma}_{SS}^{-1} \mathbf{A} \boldsymbol{\Theta}_0 \right) \\
&= \boldsymbol{\Theta}_0 + \left(\left(\boldsymbol{\Sigma}_\Theta + (\mathbf{F}\mathbf{F}^H \mathbf{A})^H (\mathbf{F}\mathbf{F}^H \boldsymbol{\Sigma}_{SS} \mathbf{F}\mathbf{F}^H)^{-1} (\mathbf{F}\mathbf{F}^H \mathbf{A}) \right)^{-1} \right. \\
&\quad \left. \left((\mathbf{F}\mathbf{F}^H \mathbf{A})^H (\mathbf{F}\mathbf{F}^H \boldsymbol{\Sigma}_{SS} \mathbf{F}\mathbf{F}^H)^{-1} (\mathbf{F}\mathbf{F}^H \mathbf{y}) - (\mathbf{F}\mathbf{F}^H \mathbf{A})^H (\mathbf{F}\mathbf{F}^H \boldsymbol{\Sigma}_{SS} \mathbf{F}\mathbf{F}^H)^{-1} (\mathbf{F}\mathbf{F}^H \mathbf{A}) \boldsymbol{\Theta}_0 \right) \right) \\
&= \boldsymbol{\Theta}_0 + \left(\boldsymbol{\Sigma}_\Theta + \tilde{\mathbf{A}}_{SS}^H \mathbf{F}^H \mathbf{F} \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \mathbf{F}^H \mathbf{F} \tilde{\mathbf{A}}_{SS}^H \right)^{-1} \left(\tilde{\mathbf{A}}_{SS}^H \mathbf{F}^H \mathbf{F} \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \mathbf{F}^H \mathbf{F} \tilde{\mathbf{y}}_{SS} - \tilde{\mathbf{A}}_{SS}^H \mathbf{F}^H \mathbf{F} \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \mathbf{F}^H \mathbf{F} \tilde{\mathbf{A}}_{SS}^H \boldsymbol{\Theta}_0 \right) \\
&= \boldsymbol{\Theta}_0 + \left(\boldsymbol{\Sigma}_\Theta + \tilde{\mathbf{A}}_{SS}^H \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \tilde{\mathbf{A}}_{SS} \right)^{-1} \left(\tilde{\mathbf{L}} - \tilde{\mathbf{A}}_{SS}^H \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \tilde{\mathbf{A}}_{SS} \boldsymbol{\Theta}_0 \right)
\end{aligned} \tag{6.40}$$

with the matrix of mean square prediction errors (see also eq. (6.8))

$$\text{MSPE}\{\tilde{\boldsymbol{\Theta}}\} = \left(\boldsymbol{\Sigma}_\Theta^{-1} + \tilde{\mathbf{A}}_{SS}^H \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \tilde{\mathbf{A}}_{SS} \right)^{-1} \tag{6.41}$$

where $\tilde{\mathbf{L}} = \tilde{\mathbf{A}}_{SS}^H \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \tilde{\mathbf{y}}_{SS}$. If $\tilde{\mathbf{a}}$ is an element of $\tilde{\mathbf{A}}$ such that

$$\tilde{\mathbf{A}}_{SS} = \begin{bmatrix} \tilde{\mathbf{a}}_0(\mathbf{f}_0) & \tilde{\mathbf{a}}_1(\mathbf{f}_0) & \cdots & \tilde{\mathbf{a}}_{m-1}(\mathbf{f}_0) \\ \tilde{\mathbf{a}}_0(\mathbf{f}_1) & \tilde{\mathbf{a}}_1(\mathbf{f}_1) & \cdots & \tilde{\mathbf{a}}_{m-1}(\mathbf{f}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{a}}_0(\mathbf{f}_{N-1}) & \tilde{\mathbf{a}}_1(\mathbf{f}_{N-1}) & \cdots & \tilde{\mathbf{a}}_{m-1}(\mathbf{f}_{N-1}) \end{bmatrix} \tag{6.42}$$

or a more compact form is

$$\tilde{\mathbf{A}}_{SS} = [\tilde{\mathbf{a}}]_{ip} = [\tilde{\mathbf{a}}_p(\mathbf{f}_i)] \text{ for } p = 0, 1, \dots, m-1 \text{ and } i = 0, 1, \dots, N-1, \tag{6.43}$$

hence, we can write

$$\tilde{\mathbf{A}}_{SS}^H \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \tilde{\mathbf{A}}_{SS} = \left[\sum_{i=0}^{N-1} \tilde{\mathbf{a}}_p^*(\mathbf{f}_i) \tilde{\boldsymbol{\Sigma}}_{SS}^{-1}(\mathbf{f}_i) \tilde{\mathbf{a}}_q(\mathbf{f}_i) \right]_{pq} \tag{6.44}$$

and

$$\tilde{\mathbf{A}}_{SS}^H \tilde{\boldsymbol{\Sigma}}_{SS}^{-1} \tilde{\mathbf{y}}_{SS} = \left[\sum_{i=0}^{N-1} \tilde{\mathbf{a}}_p^*(\mathbf{f}_i) \tilde{\boldsymbol{\Sigma}}_{SS}^{-1}(\mathbf{f}_i) \tilde{\mathbf{y}}(\mathbf{f}_i) \right]_p \text{ for } (p, q) = 0, 1, \dots, m-1 \tag{6.45}$$

In addition, with (6.44), the test statistic of (6.27) can be easily transformed using

$$\frac{\mathbf{R}}{\sigma_0^2} = \sigma_0^{-2} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0)^T \mathbf{C}_\Theta^{-1} \left(\mathbf{C}_\Theta + (\tilde{\mathbf{A}}_{SS}^H \tilde{\mathbf{C}}_{SS}^{-1} \tilde{\mathbf{A}}_{SS})^{-1} \right) \mathbf{C}_\Theta^{-1} (\tilde{\boldsymbol{\Theta}} - \boldsymbol{\Theta}_0) \tag{6.46}$$

where $\tilde{\boldsymbol{\Sigma}}_{SS} = \sigma_0^2 \tilde{\mathbf{C}}_{SS} = \sigma_0^2 \mathbf{F}^H \mathbf{C}_n \mathbf{F}$. It should be noted that if we assume sufficiently large number of data, N , or as $N \rightarrow \infty$, then the summation term in (6.39) becomes

where vectors \mathbf{y} and \mathbf{n} have dimension $KN \times 1$. The $KN \times m$ coefficient matrix \mathbf{A} consists of $N \times m$ sub-matrices \mathbf{A}_i , for $i = 0, 1, \dots, K-1$. The $KN \times KN$ cross-covariance matrix of the data is given by

$$\boldsymbol{\Sigma}_{SM} = [\boldsymbol{\Sigma}_{ij}] \text{ with } (i, j) = 0, 1, \dots, K-1 \quad (6.50)$$

where the $[i, j]$ block $\boldsymbol{\Sigma}_{ij}$ is the cross-covariance matrix between sensors i and j . Each block matrix has dimension $N \times N$ and its elements are cross-covariance (or cross-correlation) functions r_{ij} of sensors i and j

$$\boldsymbol{\Sigma}_{ij} = [r_{ij}(m-n)]_{m,n} \text{ for } (m, n) = 0, 1, \dots, N-1 \quad (6.51)$$

Since $\boldsymbol{\Sigma}_{SM}$ to be inverted is large, we transform it to a block diagonal matrix using a transformation matrix defined in (6.33) through (6.36); then only diagonal blocks are inverted. The transformation matrix is defined by

$$\mathbf{F}_{SM} = \text{diag}(\mathbf{F}_0, \mathbf{F}_1, \dots, \mathbf{F}_i, \dots, \mathbf{F}_{K-1}) \text{ with } \mathbf{F}_i = \mathbf{F} \text{ for all } i = 0, 1, \dots, N-1 \quad (6.52)$$

Then, applying (6.52) to the observation \mathbf{y}

$$\tilde{\mathbf{y}} = \mathbf{F}_{SM}^H \mathbf{y} \quad (6.53)$$

and

$$\mathbf{F}_{SM}^H \boldsymbol{\Sigma}_{SM} \mathbf{F}_{SM} = [\mathbf{F}^H \boldsymbol{\Sigma}_{ij} \mathbf{F}] \quad (6.54)$$

The $[i, j]$ block $\mathbf{F}^H \boldsymbol{\Sigma}_{ij} \mathbf{F}$ becomes a $K \times K$ diagonal matrix (see Appendix C), given by

$$\mathbf{F}^H \boldsymbol{\Sigma}_{ij} \mathbf{F} = \text{diag} \left(\sum_{k=0}^{N-1} r_{ij}(k) e^{-i2\pi k \Delta x f_0}, \sum_{k=0}^{N-1} r_{ij}(k) e^{-i2\pi k \Delta x f_1}, \dots, \sum_{k=0}^{N-1} r_{ij}(k) e^{-i2\pi k \Delta x f_{N-1}} \right) \quad (6.55)$$

Next, rearranging rows and columns of (6.53) and (6.54) [by applying the $KN \times KN$ permutation matrix, \mathbf{Z}_{SM} , constructed by the algorithm of (C.12) with $M' = K$ and $N' = N$, (see Appendix C)], we have

$$\tilde{\mathbf{y}}_{SM} = \mathbf{Z}_{SM}^H \tilde{\mathbf{y}} = \begin{bmatrix} \tilde{\mathbf{y}}(\mathbf{f}_0) \\ \tilde{\mathbf{y}}(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{y}}(\mathbf{f}_{N-1}) \end{bmatrix} = \begin{bmatrix} \tilde{y}_0(\mathbf{f}_0) \\ \tilde{y}_1(\mathbf{f}_0) \\ \vdots \\ \tilde{y}_{K-1}(\mathbf{f}_0) \\ \tilde{y}_0(\mathbf{f}_1) \\ \tilde{y}_1(\mathbf{f}_1) \\ \vdots \\ \tilde{y}_{K-1}(\mathbf{f}_1) \\ \vdots \\ \tilde{y}_0(\mathbf{f}_{N-1}) \\ \tilde{y}_1(\mathbf{f}_{N-1}) \\ \vdots \\ \tilde{y}_{K-1}(\mathbf{f}_{N-1}) \end{bmatrix} \quad (6.56)$$

with its covariance matrix

$$\begin{aligned} \tilde{\Sigma}_{SM} &= \mathbf{Z}_{SM}^H \mathbf{F}_{SM}^H \Sigma_{SM} \mathbf{F}_{SM} \mathbf{Z}_{SM} \\ &= \text{diag}(\tilde{\Sigma}_{SM}(\mathbf{f}_0), \tilde{\Sigma}_{SM}(\mathbf{f}_1), \dots, \tilde{\Sigma}_{SM}(\mathbf{f}_i), \dots, \tilde{\Sigma}_{SM}(\mathbf{f}_{N-1})) \end{aligned} \quad (6.57)$$

where the i^{th} block $\tilde{\Sigma}_{SM}(\mathbf{f}_i)$ is

$$\tilde{\Sigma}_{SM}(\mathbf{f}_i) = \begin{bmatrix} \sum_{k=0}^{N-1} r_{00}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} & \sum_{k=0}^{N-1} r_{01}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} & \dots & \sum_{k=0}^{N-1} r_{0(K-1)}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} \\ \sum_{k=0}^{N-1} r_{10}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} & \sum_{k=0}^{N-1} r_{11}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} & \dots & \sum_{k=0}^{N-1} r_{1(K-1)}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} r_{(K-1)0}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} & \sum_{k=0}^{N-1} r_{(K-1)1}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} & \dots & \sum_{k=0}^{N-1} r_{(K-1)(K-1)}(k) e^{-i2\pi k \Delta x \mathbf{f}_i} \end{bmatrix} \quad (6.58)$$

Therefore, the predicted parameters of Θ are

$$\tilde{\Theta} = \Theta_0 + \left(\Sigma_{\Theta} + \tilde{\mathbf{A}}_{SM}^H \tilde{\Sigma}_{SM}^{-1} \tilde{\mathbf{A}}_{SM} \right)^{-1} \left(\tilde{\mathbf{A}}_{SM}^H \tilde{\Sigma}_{SM}^{-1} \tilde{\mathbf{y}}_{SM} - \tilde{\mathbf{A}}_{MS}^H \tilde{\Sigma}_{SM}^{-1} \tilde{\mathbf{A}}_{SM} \Theta_0 \right) \quad (6.59)$$

where

$$\tilde{\mathbf{A}}_{SM} = \begin{bmatrix} \tilde{\mathbf{A}}_{SM}(f_0) \\ \tilde{\mathbf{A}}_{SM}(f_1) \\ \vdots \\ \tilde{\mathbf{A}}_{SM}(f_i) \\ \vdots \\ \tilde{\mathbf{A}}_{SM}(f_{N-1}) \end{bmatrix} \quad (6.60)$$

The matrix $\tilde{\mathbf{A}}_{SM}$ has dimension $KN \times m$. The $K \times m$ block $\tilde{\mathbf{A}}_{SM}(f_i)$ with its elements $\tilde{\mathbf{a}}$ is

$$\tilde{\mathbf{A}}_{SM}(f_i) = \begin{bmatrix} \tilde{a}_{00}(f_i) & \tilde{a}_{01}(f_i) & \cdots & \tilde{a}_{0(m-1)}(f_i) \\ \tilde{a}_{10}(f_i) & \tilde{a}_{11}(f_i) & \cdots & \tilde{a}_{1(m-1)}(f_i) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{(K-1)0}(f_i) & \tilde{a}_{(K-1)1}(f_i) & \cdots & \tilde{a}_{(K-1)(m-1)}(f_i) \end{bmatrix} \quad (6.61)$$

or in a compact form

$$\tilde{\mathbf{A}}_{SM}(f_i) = [\tilde{a}_{jk}(f_i)] \text{ for } j = 0, 1, \dots, K-1 \text{ and } k = 0, 1, \dots, m-1 \quad (6.62)$$

We have

$$\tilde{\mathbf{A}}_{SM}^H \tilde{\Sigma}_{SM}^{-1} \tilde{\mathbf{A}}_{SM} = \sum_{i=0}^{N-1} \tilde{\mathbf{A}}_{SM}^H(f_i) \tilde{\Sigma}_{SM}^{-1}(f_i) \tilde{\mathbf{A}}_{SM}(f_i) \quad (6.63)$$

and

$$\tilde{\mathbf{A}}_{SM}^H \tilde{\Sigma}_{SM}^{-1} \tilde{\mathbf{y}}_{SM} = \sum_{i=0}^{N-1} \tilde{\mathbf{A}}_{SM}^H(f_i) \tilde{\Sigma}_{SM}^{-1}(f_i) \tilde{\mathbf{y}}_{SM}(f_i) \quad (6.64)$$

We easily have a test statistic (not shown here) based on (6.63); the relative increase is analogous to (6.46) with subscript ‘‘SM’’ instead of ‘‘SS’’. Note that, for a large data vector, the psd of (6.58) can be approximated by

$$\tilde{\Sigma}_{SM}(f_i) \cong \frac{1}{\Delta X} \begin{bmatrix} P_{00}(f_i) & P_{01}(f_i) & \cdots & P_{0(K-1)}(f_i) \\ P_{10}(f_i) & P_{11}(f_i) & \cdots & P_{1(K-1)}(f_i) \\ \vdots & \vdots & \ddots & \vdots \\ P_{(K-1)0}(f_i) & P_{(K-1)1}(f_i) & \cdots & P_{(K-1)(K-1)}(f_i) \end{bmatrix} \quad (6.65)$$

Thus, each block has dimension $M \times M$. At frequency f_i , element $P_{mn}(f_i)$ is the auto- and cross- spectrum density of sensors m and n for $n = m$ and $n \neq m$, respectively.

6.3.3 Multiple data tracks and single sensor (MS)

For the case of multiple data tracks, we assume we have a large number of tracks, say M , and N data samples per track and track spacing is Δy . The $MN \times 1$ data vector is

$$\mathbf{y} = \begin{bmatrix} y(0,0) \\ y(1,0) \\ \vdots \\ y(N-1,0) \\ \hline y(0,1) \\ y(1,1) \\ \vdots \\ y(N-1,1) \\ \hline \vdots \\ \hline y(0,M-1) \\ y(1,M-1) \\ \vdots \\ y(N-1,M-1) \end{bmatrix} \begin{array}{l} \text{track 0} \\ \\ \\ \text{track 1} \\ \\ \\ \vdots \\ \\ \text{track } M-1 \end{array} \quad (6.66)$$

In Appendix C, the $MN \times MN$ cross-covariance matrix is given and has the following structure

$$\Sigma_{MS} = \begin{bmatrix} \Sigma_{00} & \Sigma_{01} & \cdots & \Sigma_{0j} & \cdots & \Sigma_{0(M-1)} \\ \Sigma_{10} & \Sigma_{11} & \cdots & \Sigma_{1j} & \cdots & \Sigma_{1(M-1)} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \Sigma_{i0} & \Sigma_{i1} & \cdots & \Sigma_{ij} & \cdots & \Sigma_{i(M-1)} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \Sigma_{(M-1)0} & \Sigma_{(M-1)1} & \cdots & \Sigma_{(M-1)j} & \cdots & \Sigma_{(M-1)(M-1)} \end{bmatrix} \quad (6.67)$$

where Σ_{ij} is the $N \times N$ block cross-covariance matrix between tracks i and j .

Rearranging (6.66) and (6.67), based on orthogonal transformations and permutations (see Appendix C), we finally have

$$\tilde{\mathbf{y}}_{\text{MS}} = \begin{bmatrix} \tilde{\mathbf{y}}_{\text{MS}}(\mathbf{f}_0) \\ \tilde{\mathbf{y}}_{\text{MS}}(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{y}}_{\text{MS}}(\mathbf{f}_i) \\ \vdots \\ \tilde{\mathbf{y}}_{\text{MS}}(\mathbf{f}_{N-1}) \end{bmatrix} \text{ with } \tilde{\mathbf{y}}_{\text{MS}}(\mathbf{f}_i) = \begin{bmatrix} \tilde{y}(\mathbf{f}_i, \mathbf{f}'_0) \\ \tilde{y}(\mathbf{f}_i, \mathbf{f}'_1) \\ \vdots \\ \tilde{y}(\mathbf{f}_i, \mathbf{f}'_{M-1}) \end{bmatrix} \text{ for } i = 0, 1, \dots, N-1 \quad (6.68)$$

and

$$\tilde{\Sigma}_{\text{MS}} = \text{diag}(\tilde{\Sigma}_{\text{MS}}(\mathbf{f}_0), \tilde{\Sigma}_{\text{MS}}(\mathbf{f}_1), \dots, \tilde{\Sigma}_{\text{MS}}(\mathbf{f}_i), \dots, \tilde{\Sigma}_{\text{MS}}(\mathbf{f}_{N-1})) \quad (6.69)$$

where the i^{th} block, $\tilde{\Sigma}_{\text{MS}}(\mathbf{f}_i)$, is

$$\tilde{\Sigma}_{\text{MS}}(\mathbf{f}_i) = \text{diag} \left(\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(\mathbf{k}, \mathbf{k}') e^{-i2\pi\mathbf{k}(\mathbf{k}\Delta x \mathbf{f}_i + \mathbf{k}'\Delta y \mathbf{f}'_0)}, \sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(\mathbf{k}, \mathbf{k}') e^{-i2\pi(\mathbf{f}_i \Delta x \mathbf{k} + \mathbf{f}'_i \Delta y \mathbf{k}')} , \dots, \sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(\mathbf{k}, \mathbf{k}') e^{-i2\pi(\mathbf{k}\Delta x \mathbf{f}_i + \mathbf{k}'\Delta y \mathbf{f}'_{M-1})} \right) \quad (6.70)$$

Note that only blocks $\tilde{\Sigma}_{\text{MS}}(\mathbf{f})$ are inverted. As such, the predicted parameters of Θ are

$$\tilde{\Theta} = \Theta_0 + (\Sigma_{\Theta} + \tilde{\mathbf{A}}_{\text{MS}}^H \tilde{\Sigma}_{\text{MS}}^{-1} \tilde{\mathbf{A}}_{\text{MS}})^{-1} (\tilde{\mathbf{A}}_{\text{MS}}^H \tilde{\Sigma}_{\text{MS}}^{-1} \tilde{\mathbf{y}}_{\text{MS}} - \tilde{\mathbf{A}}_{\text{MS}}^H \tilde{\Sigma}_{\text{MS}}^{-1} \tilde{\mathbf{A}}_{\text{MS}} \Theta_0) \quad (6.71)$$

where

$$\tilde{\mathbf{A}}_{\text{MS}} = \begin{bmatrix} \tilde{\mathbf{A}}_{\text{MS}}(\mathbf{f}_0) \\ \tilde{\mathbf{A}}_{\text{MS}}(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{A}}_{\text{MS}}(\mathbf{f}_i) \\ \vdots \\ \tilde{\mathbf{A}}_{\text{MS}}(\mathbf{f}_{N-1}) \end{bmatrix} \quad (6.72)$$

The matrix $\tilde{\mathbf{A}}_{\text{MS}}$ has dimension $MN \times m$. The $M \times m$ block $\tilde{\mathbf{A}}_{\text{MS}}(\mathbf{f}_i)$ with elements $\tilde{\mathbf{a}}$ is

$$\tilde{\mathbf{A}}_{\text{SM}}(\mathbf{f}_i) = \begin{bmatrix} \tilde{\mathbf{a}}_{00}(\mathbf{f}_i, \mathbf{f}'_0) & \tilde{\mathbf{a}}_{01}(\mathbf{f}_i, \mathbf{f}'_0) & \cdots & \tilde{\mathbf{a}}_{0(m-1)}(\mathbf{f}_i, \mathbf{f}'_0) \\ \tilde{\mathbf{a}}_{10}(\mathbf{f}_i, \mathbf{f}'_1) & \tilde{\mathbf{a}}_{11}(\mathbf{f}_i, \mathbf{f}'_1) & \cdots & \tilde{\mathbf{a}}_{1(m-1)}(\mathbf{f}_i, \mathbf{f}'_1) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{a}}_{(K-1)0}(\mathbf{f}_i, \mathbf{f}'_{M-1}) & \tilde{\mathbf{a}}_{(K-1)1}(\mathbf{f}_i, \mathbf{f}'_{M-1}) & \cdots & \tilde{\mathbf{a}}_{(K-1)(m-1)}(\mathbf{f}_i, \mathbf{f}'_{M-1}) \end{bmatrix} \quad (6.73)$$

or in a compact form

$$\tilde{\mathbf{A}}_{MS}(\mathbf{f}_i) = [\tilde{\mathbf{a}}_{jk}(\mathbf{f}_i, \mathbf{f}'_j)] \text{ for } j = 0, 1, \dots, M-1 \text{ and } k = 0, 1, \dots, m-1 \quad (6.74)$$

Similar to (6.63) and (6.64) but considering two dimensions, we have

$$\tilde{\mathbf{A}}_{MS}^H \tilde{\Sigma}_{MS}^{-1}(\mathbf{f}_i) \tilde{\mathbf{A}}_{MS} = \sum_{i=0}^{N-1} \tilde{\mathbf{A}}_{MS}^H(\mathbf{f}_i) \tilde{\Sigma}_{MS}^{-1}(\mathbf{f}_i) \tilde{\mathbf{A}}_{MS}(\mathbf{f}_i) \quad (6.75)$$

and

$$\tilde{\mathbf{A}}_{MS}^H \tilde{\Sigma}_{MS}^{-1} \tilde{\mathbf{y}}_{MS} = \sum_{i=0}^{N-1} \tilde{\mathbf{A}}_{MS}^H(\mathbf{f}_i) \tilde{\Sigma}_{MS}^{-1}(\mathbf{f}_i) \tilde{\mathbf{y}}_{MS}(\mathbf{f}_i) \quad (6.76)$$

Then, test statistic is easily obtained. For a large number of data, we may approximate

$$\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(k, k') e^{-i2\pi(f_i \Delta x k + f'_j \Delta y k')} \cong \frac{1}{\Delta x \Delta y} P(\mathbf{f}_i, \mathbf{f}'_j) \quad (6.77)$$

where P is the psd of gradient noise. Thus, we can approximate (6.75) and (6.76) in terms of (6.77) as

$$\left[\tilde{\mathbf{A}}_{MS}^H \tilde{\Sigma}_{MS}^{-1} \tilde{\mathbf{A}}_{MS} \right]_{pq} \cong \left[\Delta x \Delta y \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \tilde{\mathbf{a}}_{jp}^*(\mathbf{f}_i, \mathbf{f}'_j) P^{-1}(\mathbf{f}_i, \mathbf{f}'_j) \tilde{\mathbf{a}}_{jq}(\mathbf{f}_i, \mathbf{f}'_j) \right] \quad (6.78)$$

and

$$\left[\tilde{\mathbf{A}}_{MS}^H \tilde{\Sigma}_{MS}^{-1} \tilde{\mathbf{y}}_{MS} \right]_p \cong \left[\Delta x \Delta y \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \tilde{\mathbf{a}}_{jp}^*(\mathbf{f}_i, \mathbf{f}'_j) P^{-1}(\mathbf{f}_i, \mathbf{f}'_j) \tilde{\mathbf{y}}_{jp}(\mathbf{f}_i, \mathbf{f}'_j) \right] \quad (6.79)$$

for $(p, q) = 0, 1, \dots, m-1$.

6.3.4 Multiple data tracks and multiple sensors (MM)

For this case, we assume K sensors. The $KMN \times 1$ data vector is

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_i \\ \vdots \\ \mathbf{y}_{K-1} \end{bmatrix} \quad \text{with } \mathbf{y}_i = \begin{bmatrix} y_i(0,0) \\ y_i(1,0) \\ \vdots \\ y_i(N-1,0) \\ \hline y_i(0,1) \\ y_i(1,1) \\ \vdots \\ y_i(N-1,1) \\ \hline \vdots \\ \hline y_i(0,M-1) \\ y_i(1,M-1) \\ \vdots \\ y_i(N-1,M-1) \end{bmatrix} \quad \begin{array}{l} \text{track 0} \\ \\ \\ \text{track 1} \\ \\ \\ \vdots \\ \\ \text{track } M-1 \end{array} \quad \text{for } i = 0,1,\dots,K-1 \quad (6.80)$$

From Appendix C, the $KMN \times KMN$ cross-covariance matrix is given and has the following structure

$$\Sigma_{MM} = \begin{bmatrix} \Sigma_{00} & \Sigma_{01} & \cdots & \Sigma_{0j} & \cdots & \Sigma_{0(K-1)} \\ \Sigma_{10} & \Sigma_{11} & \cdots & \Sigma_{1j} & \cdots & \Sigma_{1(K-1)} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \Sigma_{i0} & \Sigma_{i1} & \cdots & \Sigma_{ij} & \cdots & \Sigma_{i(K-1)} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \Sigma_{(K-1)0} & \Sigma_{(K-1)1} & \cdots & \Sigma_{(K-1)j} & \cdots & \Sigma_{(K-1)(K-1)} \end{bmatrix} \quad (6.81)$$

The $[i,j]$ block, Σ_{ij} , having dimension $MN \times MN$, is

$$\Sigma_{ij} = \begin{bmatrix} \Sigma(0,0) & \Sigma(0,1) & \cdots & \Sigma(0,n) & \cdots & \Sigma(0,M-1) \\ \Sigma(1,0) & \Sigma(1,1) & \cdots & \Sigma(1,n) & \cdots & \Sigma(1,M-1) \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \Sigma(m,0) & \Sigma(m,1) & \cdots & \Sigma(m,n) & \cdots & \Sigma(m,M-1) \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \Sigma(M-1,0) & \Sigma(M-1,1) & \cdots & \Sigma(M-1,n) & \cdots & \Sigma(M-1,M-1) \end{bmatrix} \quad (6.82)$$

The $[n,m]$ block, $\Sigma(m,n)$, is the $N \times N$ cross-covariance matrix between tracks m and n given by

$$\Sigma(m,n) = [r_{mn}(i-j, m-n)]_{ij} \text{ for } (i,j) = 0,1,\dots,N-1 \quad (6.83)$$

where r_{mn} is a cross-correlation (covariance) function between tracks m and n .

Applying orthogonal transformations and permutations in Appendix C throughout (6.80) and (6.81), we finally have

$$\tilde{\mathbf{y}}_{MM} = \begin{bmatrix} \tilde{\mathbf{y}}_{MM}(f_0) \\ \tilde{\mathbf{y}}_{MM}(f_1) \\ \vdots \\ \tilde{\mathbf{y}}_{MM}(f_i) \\ \vdots \\ \tilde{\mathbf{y}}_{MM}(f_{N-1}) \end{bmatrix} \quad (6.84)$$

with the $KM \times 1$ vector

$$\tilde{\mathbf{y}}_{MM}(f_i) = \begin{bmatrix} \tilde{\mathbf{y}}(f_i, f'_0) \\ \tilde{\mathbf{y}}(f_i, f'_1) \\ \vdots \\ \tilde{\mathbf{y}}(f_i, f'_j) \\ \vdots \\ \tilde{\mathbf{y}}(f_i, f'_{M-1}) \end{bmatrix} \text{ for } i = 0,1,\dots,N-1 \quad (6.85)$$

and the $K \times 1$ vector

$$\tilde{\mathbf{y}}(f_i, f'_j) = \begin{bmatrix} \tilde{\mathbf{y}}_0(f_i, f'_j) \\ \tilde{\mathbf{y}}_1(f_i, f'_j) \\ \vdots \\ \tilde{\mathbf{y}}_{K-1}(f_i, f'_j) \end{bmatrix} \text{ for } j = 0,1,\dots,M-1 \quad (6.86)$$

The corresponding psd is

$$\tilde{\Sigma}_{MM} = \text{diag}(\tilde{\Sigma}_{MM}(f_0), \tilde{\Sigma}_{MM}(f_1), \dots, \tilde{\Sigma}_{MM}(f_i), \dots, \tilde{\Sigma}_{MM}(f_{N-1})) \quad (6.87)$$

The i^{th} block $\tilde{\Sigma}_{MM}(f_i)$ having dimension $KM \times KM$ is

$$\tilde{\Sigma}_{MM}(\mathbf{f}_i) = \text{diag}(\tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_0), \tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_1), \dots, \tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_j), \dots, \tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_{M-1})) \quad (6.88)$$

with the $K \times K$ block

$$\tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_j) = \left[\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r_{mn}(\mathbf{k}, \mathbf{k}') e^{-i2\pi(\Delta x k f_i + \Delta y k' f'_j)} \right]_{mn} \quad \text{for } (m, n) = 0, 1, \dots, K-1 \quad (6.89)$$

where r_{mn} is the cross-correlation function between sensors m and n (also see (6.83)).

Therefore, the predicted parameters of Θ are

$$\tilde{\Theta} = \Theta_0 + \left(\Sigma_{\Theta} + \tilde{\mathbf{A}}_{MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{A}}_{MM} \right)^{-1} \left(\tilde{\mathbf{A}}_{MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{y}}_{MM} - \tilde{\mathbf{A}}_{MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{A}}_{MM} \Theta_0 \right) \quad (6.90)$$

where

$$\tilde{\mathbf{A}}_{MM} = \begin{bmatrix} \tilde{\mathbf{A}}_{MM}(\mathbf{f}_0) \\ \tilde{\mathbf{A}}_{MM}(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{A}}_{MM}(\mathbf{f}_i) \\ \vdots \\ \tilde{\mathbf{A}}_{MM}(\mathbf{f}_{N-1}) \end{bmatrix} \quad (6.91)$$

The matrix $\tilde{\mathbf{A}}_{MM}$ has dimension $KMN \times m$. The $KM \times m$ block $\tilde{\mathbf{A}}_{MM}(\mathbf{f}_i)$ is

$$\tilde{\mathbf{A}}_{MM}(\mathbf{f}_i) = \begin{bmatrix} \tilde{\mathbf{A}}(\mathbf{f}_i, \mathbf{f}'_0) \\ \tilde{\mathbf{A}}(\mathbf{f}_i, \mathbf{f}'_1) \\ \vdots \\ \tilde{\mathbf{A}}(\mathbf{f}_i, \mathbf{f}'_j) \\ \vdots \\ \tilde{\mathbf{A}}(\mathbf{f}_i, \mathbf{f}'_{M-1}) \end{bmatrix} \quad \text{for } j = 0, 1, \dots, M-1 \quad (6.92)$$

where the $K \times m$ sub-block matrix $\tilde{\mathbf{A}}(\mathbf{f}_i, \mathbf{f}'_j)$ with elements \tilde{a} is

$$\tilde{\mathbf{A}}(\mathbf{f}_i, \mathbf{f}'_j) = \begin{bmatrix} \tilde{a}_{00}(\mathbf{f}_i, \mathbf{f}'_j) & \tilde{a}_{01}(\mathbf{f}_i, \mathbf{f}'_j) & \cdots & \tilde{a}_{0(m-1)}(\mathbf{f}_i, \mathbf{f}'_j) \\ \tilde{a}_{10}(\mathbf{f}_i, \mathbf{f}'_j) & \tilde{a}_{11}(\mathbf{f}_i, \mathbf{f}'_j) & \cdots & \tilde{a}_{1(m-1)}(\mathbf{f}_i, \mathbf{f}'_j) \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{a}_{(K-1)0}(\mathbf{f}_i, \mathbf{f}'_j) & \tilde{a}_{(K-1)1}(\mathbf{f}_i, \mathbf{f}'_j) & \cdots & \tilde{a}_{(K-1)(m-1)}(\mathbf{f}_i, \mathbf{f}'_j) \end{bmatrix} \quad (6.93)$$

From (6.91) to (6.93), we have

$$\begin{aligned}\tilde{\mathbf{A}}_{MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{A}}_{MM} &= \sum_{i=0}^{N-1} \tilde{\mathbf{A}}_{MM}^H(\mathbf{f}_i) \tilde{\Sigma}_{MM}^{-1}(\mathbf{f}_i) \tilde{\mathbf{A}}_{MM}(\mathbf{f}_i) \\ &= \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \tilde{\mathbf{A}}^H(\mathbf{f}_i, \mathbf{f}'_j) \tilde{\Sigma}^{-1}(\mathbf{f}_i, \mathbf{f}'_j) \tilde{\mathbf{A}}(\mathbf{f}_i, \mathbf{f}'_j)\end{aligned}\quad (6.94)$$

and

$$\begin{aligned}\tilde{\mathbf{A}}_{MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{y}}_{MM} &= \sum_{i=0}^{N-1} \tilde{\mathbf{A}}_{MM}^H(\mathbf{f}_i) \tilde{\Sigma}_{MM}^{-1}(\mathbf{f}_i) \tilde{\mathbf{y}}_{MM}(\mathbf{f}_i) \\ &= \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \tilde{\mathbf{A}}^H(\mathbf{f}_i, \mathbf{f}'_j) \tilde{\Sigma}^{-1}(\mathbf{f}_i, \mathbf{f}'_j) \tilde{\mathbf{y}}(\mathbf{f}_i, \mathbf{f}'_j)\end{aligned}\quad (6.95)$$

In addition, the test statistic can be easily obtained. For a large number of data, we may approximate

$$\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} \Gamma_{pq}(\mathbf{k}, \mathbf{k}') e^{-i2\pi(\mathbf{f}_i \Delta \mathbf{x} \mathbf{k} + \mathbf{f}'_j \Delta \mathbf{y} \mathbf{k}')} \cong \frac{1}{\Delta \mathbf{x} \Delta \mathbf{y}} \mathbf{P}_{pq}(\mathbf{f}_i, \mathbf{f}'_j) \quad (6.96)$$

where \mathbf{P} is the cross psd of gradient noise between sensors p and q . Thus, the psd $\tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_j)$ can be approximated:

$$\tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_j) \cong \frac{1}{\Delta \mathbf{x} \Delta \mathbf{y}} \begin{bmatrix} \mathbf{P}_{00}(\mathbf{f}_i, \mathbf{f}'_j) & \mathbf{P}_{01}(\mathbf{f}_i, \mathbf{f}'_j) & \cdots & \mathbf{P}_{0(K-1)}(\mathbf{f}_i, \mathbf{f}'_j) \\ \mathbf{P}_{10}(\mathbf{f}_i, \mathbf{f}'_j) & \mathbf{P}_{11}(\mathbf{f}_i, \mathbf{f}'_j) & \cdots & \mathbf{P}_{1(K-1)}(\mathbf{f}_i, \mathbf{f}'_j) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{P}_{(K-1)0}(\mathbf{f}_i, \mathbf{f}'_j) & \mathbf{P}_{(K-1)1}(\mathbf{f}_i, \mathbf{f}'_j) & \cdots & \mathbf{P}_{(K-1)(K-1)}(\mathbf{f}_i, \mathbf{f}'_j) \end{bmatrix} \quad (6.97)$$

CHAPTER 7

NUMERICAL DISCUSSIONS ON PARAMETER ESTIMATION

7.1 Introduction

In this section, numerical investigations for cases of multiple tracks with single and multiple sensors are conducted. Provided that matched filters give the grid location of a detected anomaly, that is the location of the highest peak in the output, and the approximated depth and orientation are acquired during the detection attempts, the estimation approach can perform as a tool of refinement of the location, orientation, and density contrast. For our study, we only show the robustness of the random effects model and its hypothesis testing with respect to the a priori information achieved by matched filtering. We thus consider the 2m x 2m x 1000m anomaly at 50m depth with the azimuthal orientations, β' , of 90°, 47°, 45°, and 0°. We evaluate the capability of estimating parameters as well as hypothesis testing in various noise environments. All least-squares adjustment examples discussed here are based on the set of simulated observations described in Chapter 5. To provide intuitive discussions of (1) the numerical results of parameter estimation and (2) the influence of noise levels affecting the estimation and hypothesis testing, we demonstrate the estimation with respect to several types of observations obtained from the gradient signals due to the anomaly in different noise backgrounds such as 0.01E²/Hz and 1E²/Hz white noise alone, and the full geological background plus 1E²/Hz white noise before and after the subtraction of the geological model defined in Chapter 5. Table 7.1 lists the root-mean-squares (rms's) of the noises and the gradient signals due to the anomaly at $\beta' = 0^\circ, 45^\circ, 47^\circ, \text{ and } 90^\circ$.

For simplicity, we assume we have many data with 30m sampling interval (i.e., 3s temporal spacing with the helicopter's speed of 10m/s) and 30m track spacing. The psd of the geological noise backgrounds is computed according to (6.97) using Psd models 1 or 2. To avoid the confusion of terminology used in this study, the terms “estimate”, “predict”, and “adjust” are used interchangeably unless otherwise specified.

Since we linearize the observation model, iterations are required. The equation of (6.4) refers to the uncorrelated-ness between the observations and the a priori information, where $\Sigma_\Theta = \sigma_0^2 C_\Theta$ is guessed from the stage of signal detection attempts as a covariance matrix of Θ . We consider updates to the predicted parameters while keeping C_Θ unchanged as follows (e.g., the case of multiple data tracks and multiple sensors, MM):

$$\tilde{\Theta}_i = \tilde{\Theta}_{i-1} + \left(\Sigma_\Theta^{-1} + \tilde{\mathbf{A}}_{i,MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{A}}_{i,MM} \right)^{-1} \left(\tilde{\mathbf{A}}_{i,MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{y}}_{MM} - \tilde{\mathbf{A}}_{i,MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{A}}_{i,MM} \tilde{\Theta}_{i-1} \right)$$

with

Grad.		rms (Eötvös)					
		Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
Signal	90°	5.347E-3	0.844E-3	5.147E-3	0.831E-3	5.100E-3	1.184E-3
	47°	5.289E-3	2.614E-3	2.930E-3	2.506E-3	3.856E-3	3.621E-3
	45°	5.291E-3	2.773E-3	2.773E-3	2.512E-3	3.739E-3	3.739E-3
	0°	5.347E-3	5.147E-3	0.844E-3	0.831E-3	1.184E-3	5.100E-3
Noise	0.01E ² /Hz	5.769E-2	5.773E-2	5.771E-2	5.771E-2	5.766E-2	5.770E-2
	1E ² /Hz	5.785E-1	5.775E-1	5.777E-1	5.775E-1	5.769E-1	5.770E-1
	WOB	5.720E+0	5.420E+0	3.715E+0	3.990E+0	3.476E+0	2.636E+0
	WB	1.218E+1	6.889E+0	7.186E+0	4.005E+0	4.119E+0	3.049E+0
$\text{rms} = \sqrt{\sum_{i=0}^{n-1} s_i^2} / n \text{ where } s_i = \text{signal or noise and } n = 400 \times 400^1$ <p>WB = full geological background plus 1E²/Hz white noise WOB = full geological background plus 1E²/Hz white noise minus a geological model²</p>							

Table 7.1: CASE I: The root-mean-square (rms) of the gradient signals due to the 2m x 2m x 1000m anomaly at the azimuthal orientations of 90°, 47°, 45° and 0° at 50m depth versus noise backgrounds.

$$\text{MSPE}\{\tilde{\Theta}_i\} = \left(\Sigma_{\Theta}^{-1} + \tilde{\mathbf{A}}_{i,MM}^H \tilde{\Sigma}_{MM}^{-1} \tilde{\mathbf{A}}_{i,MM} \right)^{-1} = \hat{\sigma}_{0,i}^2 \left(\mathbf{C}_{\Theta}^{-1} + \tilde{\mathbf{A}}_{i,MM}^H \tilde{\mathbf{C}}_{MM}^{-1} \tilde{\mathbf{A}}_{i,MM} \right)^{-1} \quad (7.2)$$

where the subscript “i” is the number of the iteration and the approximated θ_0 is obtained from the matched filter detection output. The $\tilde{\theta}_{i-1}$ are updated by the computed $\tilde{\Theta}_{i-1}$, then $\tilde{\theta}_i$ can be achieved. The $\hat{\sigma}_{0,i}^2$ is computed using eq. (6.11). Theoretically, for “i” iterations, if the adjustment properly converges, $\tilde{\Theta}_i$ should converge to zero according to the linearized equation of (6.3). The $\tilde{\Theta}_i$ should be close or equal to Θ_0 (with uncertainty implied by the a priori variance) in the statistical sense according to (6.24). Therefore, the null hypothesis H_0 is accepted if the test statistic of (6.27) is below the threshold ψ_0 for a given α .

¹ Due to the memory limitation of the Fortran 90 program used, we consider only a 12km x 12km study area (or 400 x 400 grid points with 30m spacing along track and 30m track spacing).

² A geological model is constructed using the 120m x 120m DEM (refer to Section 5.1 on page 42).

Suppose the true location of the center of the anomaly is at N21013m and E18018m, and 50m depth, and the anomaly is rotated by the different azimuthal orientations $\beta' = \{0^\circ, 45^\circ, 47^\circ, \text{ and } 90^\circ\}$. The location where the output of the matched filter is highest, is given by

$$\theta_0 = \begin{bmatrix} \text{North} \\ \text{East} \\ \text{Depth} \\ \text{Orient} \end{bmatrix} = \begin{bmatrix} 21030\text{m} \\ 18000\text{m} \\ 60\text{m} \\ \beta' \end{bmatrix} \quad (7.3)$$

where the depth of 60m is chosen in the detection process and β' also is obtained from the banked-filter process. For a strict constraint, we set $\Sigma_\Theta \rightarrow 0$, which implies $\Theta_0 = \Theta$ and that the parameters Θ are assumed known perfectly. In general, we consider Θ as “random effects” and choose the suitable values for Σ_Θ . It should be kept in mind that the smaller the covariance Σ_Θ is, the closer we gain perfect satisfaction of our strict constraint. We want to choose the a priori variances applicable to our objective of not only improved localization but also hypothesis testing. For the first example (i.e., CASE I), three possible a priori depth variances are selected to show how the adjusted solutions and their respective test statistics behave and then we shall select the most reliable one for further analyses:

$$\Sigma_\Theta = \sigma_0^2 \begin{bmatrix} \sigma_N^2 & 0 & 0 & 0 \\ 0 & \sigma_E^2 & 0 & 0 \\ 0 & 0 & \sigma_D^2 & 0 \\ 0 & 0 & 0 & \sigma_{\beta'}^2 \end{bmatrix} \quad \begin{array}{l} \sigma_N = \pm 30.00\text{m} \\ \text{with } \sigma_E = \pm 30.00\text{m} \\ \sigma_{\beta'} = \pm 3^\circ \end{array} \quad (7.4)$$

(a) $\sigma_D = \pm 0.10\text{m}$, (b) $\sigma_D = \pm 1.00\text{m}$, and (c) $\sigma_D = \pm 10.00\text{m}$

The σ_N and σ_E are given on the basis of 30m grid points which are used for the multiple sets of matched filtering detection attempts. The standard deviation of the orientation, $\sigma_{\beta'}$, is given as a reasonable guess, but could also be based on results from applying a bank of matched filters. What follow are several possible cases to test the method of parameter estimation with respect to various noise environments.

7.2 CASE I: The impact of the a priori depth variance on localization and test statistic

Prior to showing the numerical solutions of parameter estimation with respect to orientations and several types of noise backgrounds, let the impact of the a priori covariance matrix on the (depth) parameter estimates and their respective test statistics be discussed first. In this example, consider the problem of estimating the location of the 2m x 2m x 1000m anomaly, where the azimuthal orientation $\beta' = 90^\circ$ is assumed known perfectly (i.e., $\sigma_{\beta'} = 0^\circ$) and no orientation estimation is taken into account. We consider the case of multiple data tracks and the single sensor (MS) for the gradient component Γ_{33} only. The gradient signals due to the anomaly are corrupted by $0.01E^2/\text{Hz}$ white Gaussian noise (equivalently, the variance, σ_r^2 , is $1/300 E^2$); the flat psd of (6.77) is $3E^2/(\text{cyc}/\text{m})^2$ for all frequencies. We initially set Θ_0 equal to zero (refer to eq. (7.1)). For the different values of σ_D in (7.4), we assess the adjusted solutions using five iterations. These solutions are summarized in Table 7.2.

Note that selecting $\sigma_D = \pm 0.10\text{m}$ yields the estimates of the depth with slower convergence compared to the other cases. As the σ_D increases, the errors get smaller with further iterations, particularly, in the depth estimation. The fifth iterated solutions for cases (b) and (c) yield estimates of north and east position with sub-meter and 10-m errors, respectively. Both cases yield similar results in the estimates. However, the test statistic differs significantly. As such, it indicates that the choices of σ_D affect the test statistic (6.27).

All cases of σ_D^2 show reasonable convergence of the iterative solutions; the errors generally decrease with respect to the number of iterations. Note that the larger values of the assigned variance σ_D^2 affect the test statistic obtained from (6.27). For instance, the test statistic in case (c) is smallest as compared to the other cases. (It should be kept in mind also that, based on (7.1), the algorithm updates the initial approximate θ_0 and in turn Θ_0 for each iteration, while C_θ is unchanged.) A suitable choice of σ_D^2 should be carefully defined. As regards to the numerical experiments in Table 7.2, we have selected $\sigma_D^2 = 1\text{m}^2$ as the a priori variance of depth for our study because it yields reasonable convergence of the iterative solutions (particularly in the depth estimation) associated with the test statistic (6.27). However, we shall see later in CASE III that selecting $\sigma_D^2 = 100\text{m}^2$ yields more stable solutions (although the test statistic seems to be suspect with this choice).

Iteration	ΔN (m)			ΔE (m)			ΔD (m)			Test statistic, T (eq. 6.27)		
	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
0	4.795	4.795	4.795	2.546	2.546	2.546	9.983	8.954	7.418	1.7832	1.1854	0.2927
1	3.171	2.859	2.420	12.298	12.251	11.787	9.898	4.593	0.468	24.3461	9.4588	0.6142
2	1.041	0.262	0.230	8.039	8.802	8.955	9.742	0.138	0.204	26.8938	0.1159	0.2388
3	1.290	0.553	0.511	10.258	8.872	8.805	9.514	1.640	0.155	26.7164	4.0497	0.0052
4	0.017	0.493	0.481	9.149	8.831	8.986	9.216	1.274	0.137	25.6626	2.4410	0.0001
5	0.699	0.493	0.479	9.418	8.894	8.840	8.849	0.321	0.140	24.0367	0.1854	0.0000

ΔN = the absolute error of north coordinate
 ΔE = the absolute error of east coordinate
 ΔD = the absolute error of depth
(a) $\sigma_D = \pm 0.10\text{m}$; (b) $\sigma_D = \pm 1.00\text{m}$; (c) $\sigma_D = \pm 10.00\text{m}$

Table 7.2: The impact of the a priori depth variance on the parameter estimation and test statistic using Γ_{33} only.

7.3 CASE II: Predicted location and fixed orientation using individual gradient components in $0.01E^2/\text{Hz}$ (or $3E^2/(\text{cyc/m})^2$) white noise environment

We consider the case of multiple data tracks and single sensors. The gradient signals of the anomaly are corrupted by $0.01E^2/\text{Hz}$ white Gaussian noise; the flat psd of (6.77) is $3E^2/(\text{cyc/m})^2$ for all frequencies. The initial value of Θ_0 is equal to zero and the a priori covariance matrix of Θ is taken from (7.4b). The azimuthal orientation is assumed to be known completely (i.e., no orientation estimation). The parameter estimation is independently conducted for each gradient component at $\beta' = 90^\circ$, 45° , and 0° . Through 50 iterations, the absolute errors, $|\tilde{\Theta} - \Theta|$, of the location (north and east coordinates and depth) using individual gradients are plotted in Figures 7.1 to 7.3 as well as their respective test statistics.

Figure 7.1a shows the case where the anomaly is at $\beta' = 90^\circ$. The location errors vary in the range of 1 to 20 meters through the iteration process. The components Γ_{33} , Γ_{11} , and Γ_{13} yield smaller errors (a few meter level) in the north direction than the other components whose errors fall between 5 and 15 meters. The depth errors for the components Γ_{33} , Γ_{11} , and Γ_{13} are about 1 meter, while they vary in the range of 2 – 10m after 50 iterations for the other components.

Figure 7.2a, for $\beta' = 45^\circ$, shows an improvement in estimating the east coordinate using Γ_{33} , relative to the case of $\beta' = 90^\circ$. The east error is about 4m. The different orientation also affects the estimation of the north coordinate. For instance, the estimation using Γ_{11} shows an increase to 5m in the north error from 1m in the $\beta' = 90^\circ$ case, see also Figure 7.1a. The depth error varies in the range of 1 to 3m for all components.

Consider $\beta' = 0^\circ$, where the 1000m length-side of the anomaly is parallel to the north direction. This case is opposite to the case where the orientation is 90° . The error in the estimate of the north coordinate is 2m larger than that of the east coordinates using Γ_{33} , shown in Figure 7.3a. The depth errors by using Γ_{12} , Γ_{11} , and Γ_{13} oscillate like they do when using Γ_{12} , Γ_{22} , and Γ_{23} for $\beta' = 90^\circ$, shown in Figure 7.1a. These results are due to a high weight (the inverse of $\sigma_D^2 = 1\text{m}^2$) for the depth estimate. We shall see later in CASE III that taking a lower weight (i.e., the inverse of $\sigma_D^2 = 100\text{m}^2$) yields more stable solutions.

The MSPE³ (7.2) given in the equations (7.5) and (7.7) show very small correlation among the estimates using Γ_{33} where the anomaly is at $\beta' = 90^\circ$ and 0° respectively. At $\beta' = 45^\circ$, the MSPE in equation (7.6) indicates a high correlation between north and east coordinates. (For the other gradient components, the results are similar (not shown

³ In the view point of “pseudo-observation” constraints, these results turn out to be the standard deviations of the estimates of the “non-random” parameters, whereas the standard deviations of the “random effects” parameters for our case are known a priori (refer to eq. (7.4)). For further details, see Leick (1995, pp. 126) and Mikhail and Ackerman (1976, pp. 343-352).

here).) However, the adjusted horizontal coordinates are independent of the depth estimation. Examples of $\text{MSPE}\{\tilde{\Theta}_{50}\}$ at iteration 50 for Γ_{33} obtained from eq. (7.2) are:

I: The azimuthal orientation is 90° and $\hat{\sigma}_0^2 = 0.998316$

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 0.998316 \begin{bmatrix} 0.710125 & 0.000016 & 0.000000 \\ 0.000016 & 24.957517 & 0.000000 \\ 0.000000 & 0.000000 & 0.408823 \end{bmatrix} [\text{m}]^2 \quad (7.5)$$

II: The azimuthal orientation is 45° and $\hat{\sigma}_0^2 = 0.998331$

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 0.998331 \begin{bmatrix} 12.271421 & 11.622540 & 0.000000 \\ 11.622539 & 12.270313 & 0.000000 \\ 0.000000 & 0.000000 & 0.387288 \end{bmatrix} [\text{m}]^2 \quad (7.6)$$

III: The azimuthal orientation is 0° and $\hat{\sigma}_0^2 = 0.998314$

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 0.998314 \begin{bmatrix} 22.390881 & -0.000010 & 0.000000 \\ -0.000010 & 0.618668 & 0.000000 \\ 0.000000 & 0.000000 & 0.376107 \end{bmatrix} [\text{m}]^2 \quad (7.7)$$

Finally, Table 7.3 lists the root-mean-square-error (rmse) of the estimates for each gradient component, obtained from the square root of the diagonal elements of eqs. (7.2); for instance, for $\beta' = 90^\circ$ and Γ_{33} in eq. (7.5), $\text{rmse} = \sqrt{0.998316 \times 0.710125} = 0.824\text{m}$. It should be noted that the rmse's of the estimates reflect the actual estimate errors. For instance, in the case of $\beta' = 0^\circ$, the rmse of the east coordinate using Γ_{33} is 0.786m which reflect the actual east error in the magnitude of about 0.700m , illustrated in Figure 7.3a.

To assess the estimation performance by the random effects model, the test statistic of (6.27) is computed. The estimated locations (i.e., $\tilde{\Theta}$) and the a priori information (i.e., Θ_0) do not have to be numerically equal but under the null hypothesis they should be statistically equal in the sense of (6.24) with (6.26). In Figures 7.1b -7.3b, with the significant level $\alpha = 0.05$, using Γ_{33} for instance, the values of test statistic exceed the threshold ψ_0 for the first five iterations. Such a failure indicates a (significantly) large value of $\tilde{\Theta}$ relative to Θ_0 . For instance, Figure 7.1a shows the errors using Γ_{33} are about 10m at most and become stable after 5 iterations. Correspondingly, the values of the test statistic (red solid line) are significantly large, but then quickly diminish after 5 iterations.

Although there are large variations in location errors using Γ_{22} , Γ_{12} , and Γ_{23} at $\beta' = 90^\circ$ (see also Γ_{11} , Γ_{12} , and Γ_{13} at $\beta' = 0^\circ$), the test statistic is not sensitive to these errors,

	The root-mean-square-error (m)					
	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
The azimuthal orientation fixed to 90°						
Northing	0.842	14.862	0.881	5.431	0.716	6.399
Easting	4.991	9.398	8.251	8.909	6.271	6.421
Depth	0.639	0.931	0.659	0.978	0.580	0.977
The azimuthal orientation fixed to 45°						
Northing	3.500	5.688	5.504	9.435	4.882	5.043
Easting	3.499	5.184	6.055	9.435	5.135	4.797
Depth	0.622	0.826	0.846	0.867	0.756	0.748
The azimuthal orientation fixed to 0°						
Northing	4.728	8.236	7.913	15.215	7.047	7.064
Easting	0.786	0.876	12.819	10.624	7.056	0.838
Depth	0.613	0.657	0.990	0.994	0.980	0.640

Table 7.3: CASE II: The root-mean-square-error (rmse) of the adjusted parameters after 50 iterations

i.e., the test statistic is below the threshold. The null hypothesis H_0 is accepted for all iterated solutions. It thus interprets that $\tilde{\Theta}$ and Θ_0 are not statistically different; there is a consistency between the a priori information and the observations. The following discussions concern the test statistic when the signal is small in a strong noise background.

Consider Γ_{12} at iteration 50 for the case of $\beta' = 90^\circ$, where the location errors, shown in Figure 7.1a, are relatively large, compared with the errors using the other components. However, the plot in Figure 7.1b shows the values of the test statistic are smaller than the threshold (red solid line). Table 7.4 shows $A = 0.2322$, $B = 0.9993$, and the test statistic $T = 0.23$ is below $\psi_0 = 2.60$ after 50 iterations. Since the Γ_{12} signal is small relative to the noise background (i.e., $0.01E^2/\text{Hz}$), the test statistic may not provide useful information for the test.

Table 7.5 shows the ratio of the rms of the signal and the rms of the noise (SNR), taken from Table 7.1. For instance, for $\beta' = 90^\circ$, the SNR of Γ_{12} ($=0.014$) is about 17% of the SNR of Γ_{33} ($=0.093$). The test statistic for Γ_{33} is able to indicate large location errors, i.e., $T > \psi_0$ for the first 5 iterations whereas the test for Γ_{12} is not affected by large

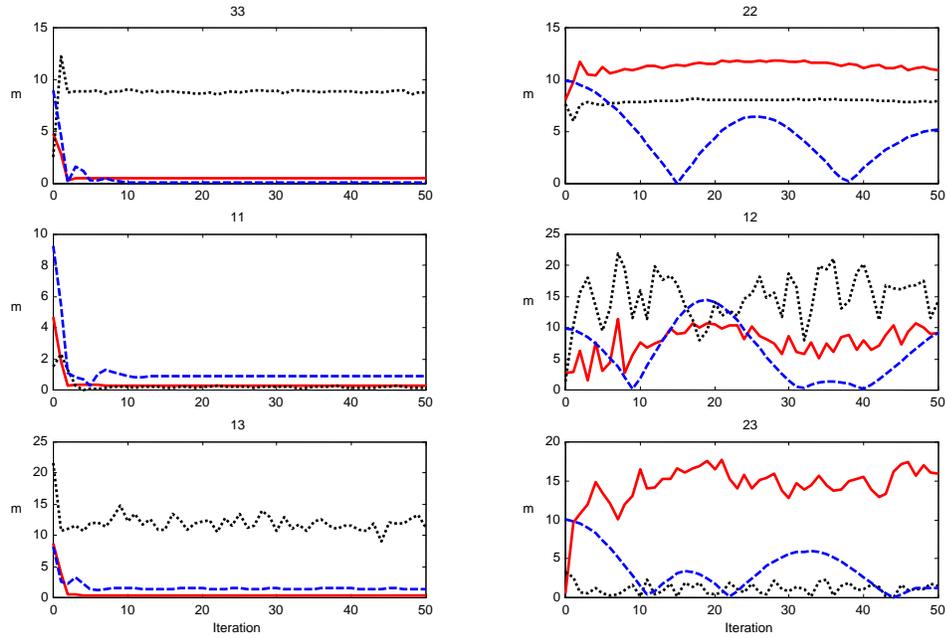
location errors, i.e., $T < \psi_0$, shown in Figure 7.1b. Interestingly, the increase of Γ_{12} at $\beta' = 45^\circ$, where $\text{SNR} = 0.044$, yields larger values of the test statistic which is consistent with large location errors. Figure 7.2b shows that the values of test statistic exceed the threshold for the first 4 iterations where the location errors significantly change. For these iterations, the hypothesis test would indicate the rejection of H_0 for the given $\alpha = 0.05$.

The values of SNR imply that the robustness of the estimation method is downgraded if the anomaly's signal is small and almost entirely corrupted by strong noise, thus making the test statistic unreliable; Γ_{12} at $\beta' = 90^\circ$ or 0° is such a case.

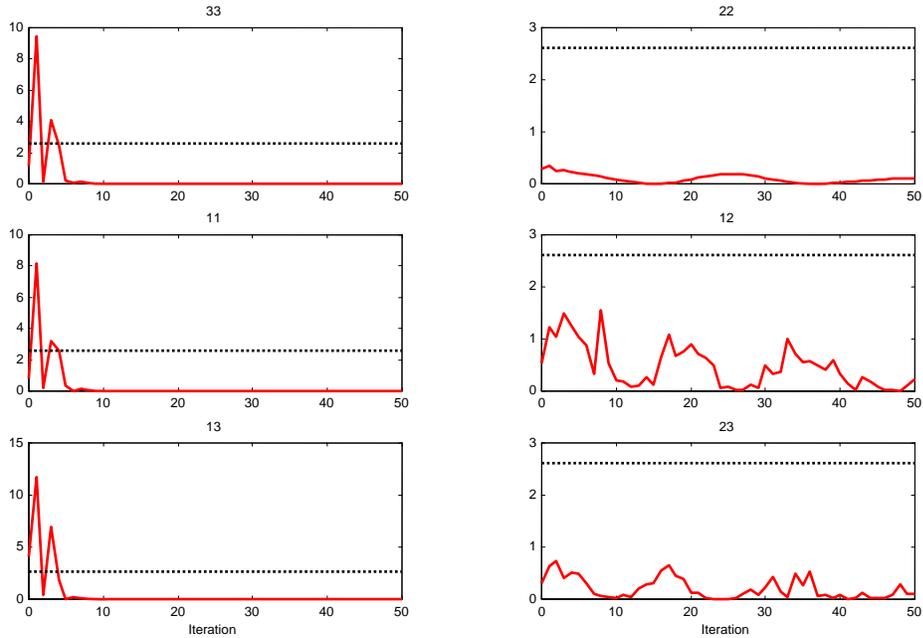
Therefore, the unwanted part of the observations (i.e., noise backgrounds) should be removed as much as possible. Alternatively, the signals should be significantly larger than the noise background. To improve the capability of the test statistic (i.e. eq. (6.27) in Table 7.4), we require a reduction of the noise background in the observation \mathbf{y} in eq. (6.13) so as to get B smaller and/or an increase of signal gradient so as to get A larger.

Table 7.6 shows the example of the test statistic operations for Γ_{12} associated with $\beta' = 90^\circ$ versus 45° (where $\text{SNR} = 0.014$ and 0.044 , respectively, in Table 7.5). It should be noted that the values of \mathbf{R}/m of the test statistic for $\beta' = 90^\circ$ are smaller than those for the case of $\beta' = 45^\circ$. The numerical analyses indicate that not only the strong noise background but also the magnitude and orientation (i.e., the signal strength) of the target anomaly greatly affect the test statistic.

Referring to these numerical results, overall, the parameter estimation and hypothesis testing using the component Γ_{33} is superior to other components with respect to $\beta' = 90^\circ$, 45° and 0° . The next section will show how a stronger noise background will affect the parameter estimation and test statistic.

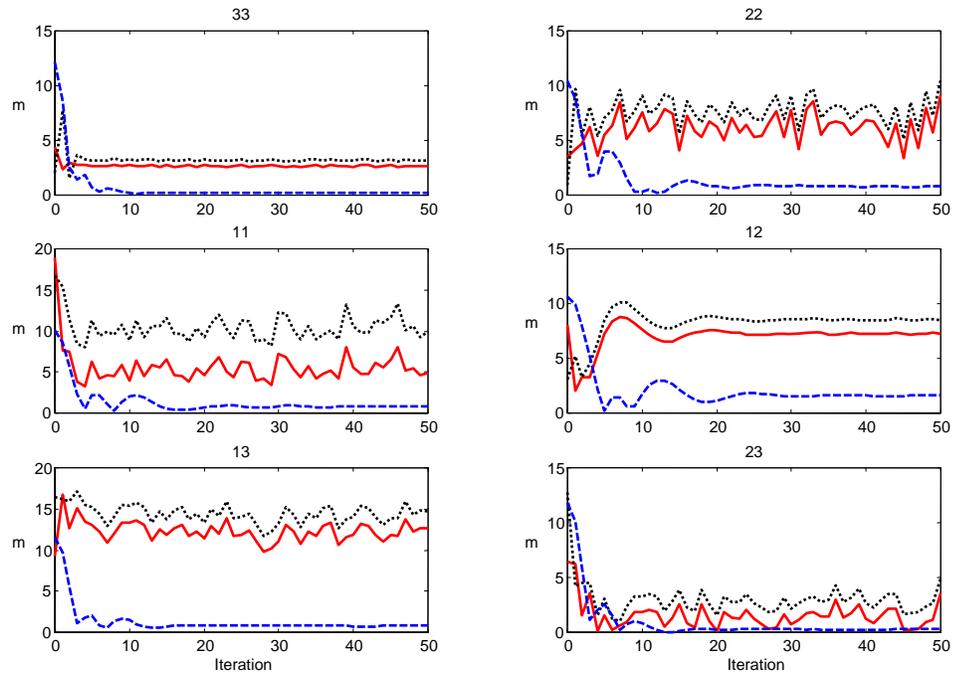


(a)

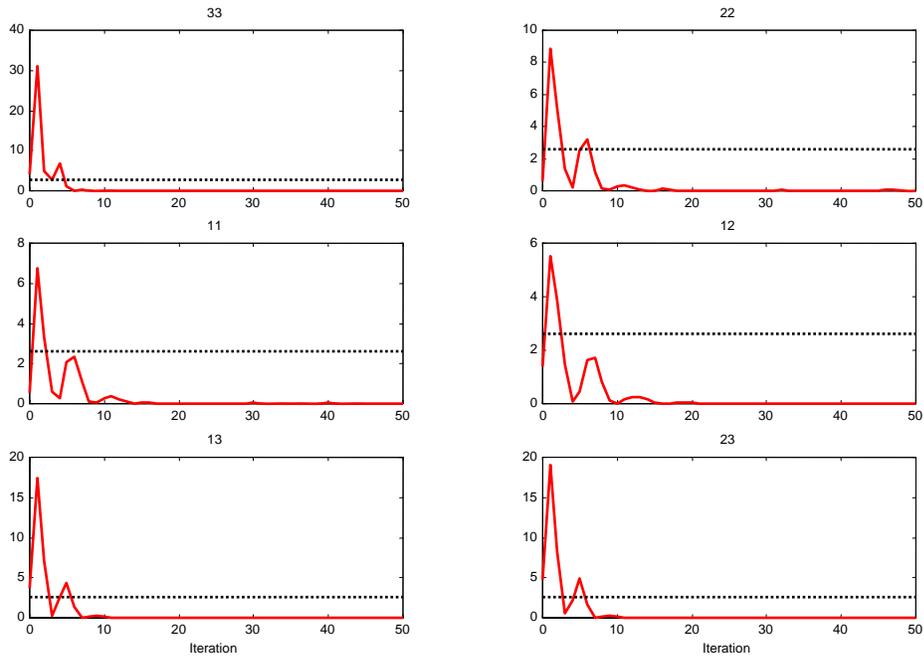


(b)

Figure 7.1: CASE II: (a) The absolute errors of north (red solid line), east (black dotted line), and depth (blue dashed line) at the azimuthal orientation of 90° ; (b) Test statistic (red solid line) and $\psi_0 = 2.60$ (black dotted line) with $m = 3$ and $n = 400 \times 400$.

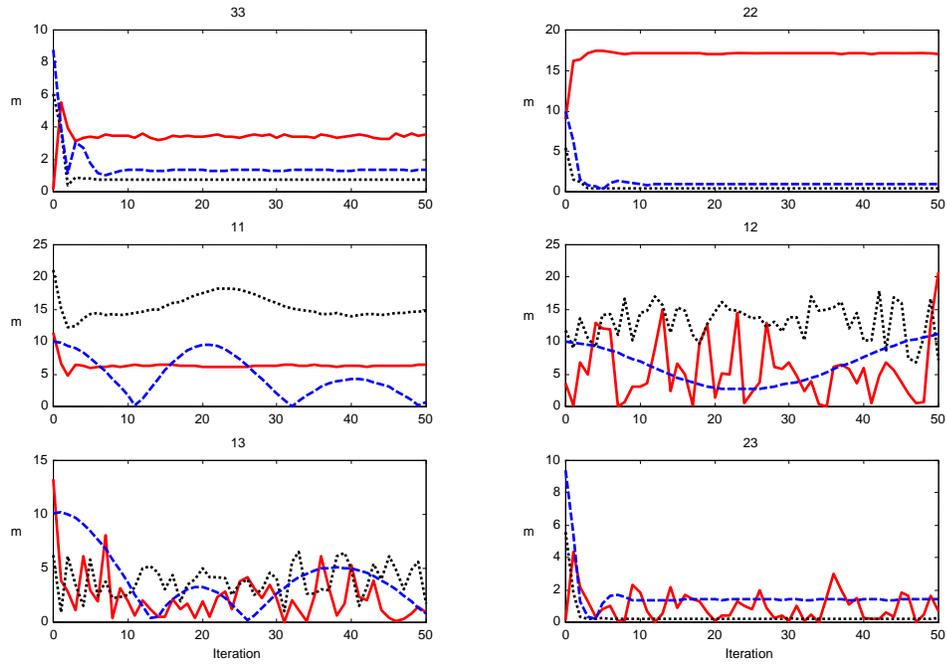


(a)

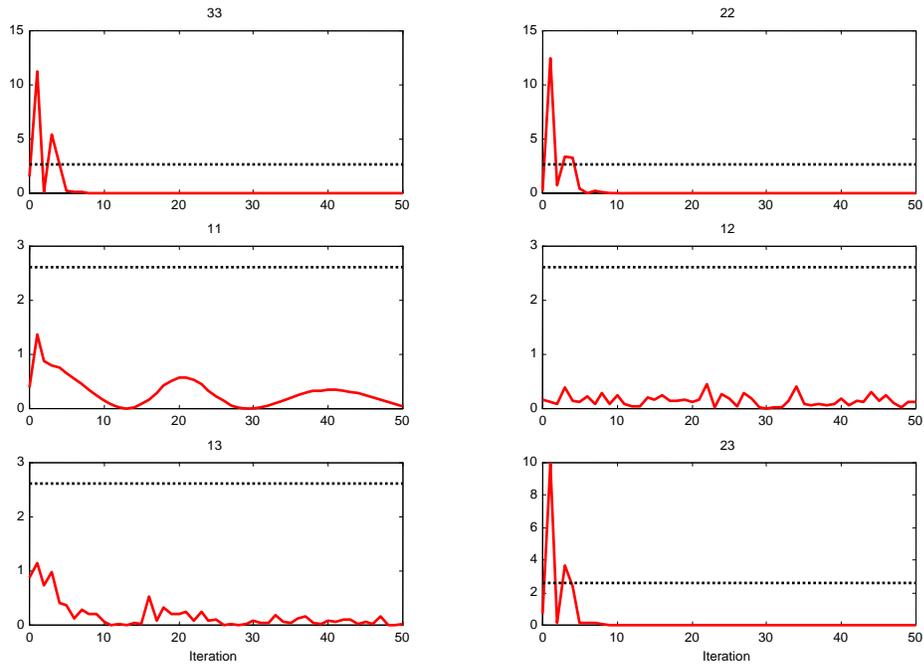


(b)

Figure 7.2: CASE II: (a) The absolute errors of north (red solid line), east (black dotted line), and depth (blue dashed line) at the azimuthal orientation of 45° ; (b) Test statistic (red solid line) and $\psi_0 = 2.60$ (black dotted line) with $m = 3$ and $n = 400 \times 400$.



(a)



(b)

Figure 7.3: CASE II: (a) The absolute errors of north (red solid line), east (black dotted line), and depth (blue dashed line) at the azimuthal orientation of 0° ; (b) Test statistic (red solid line) and $\psi_0 = 2.60$ (black dotted line) with $m = 3$ and $n = 400 \times 400$.

Grad.	Test statistic								
	90°			45°			0°		
	A	B	T	A	B	T	A	B	T
Γ_{33}	0.000038614	0.998335258	0.000038678	0.000018235	0.998349319	0.000018265	0.000022515	0.998333284	0.000022552
Γ_{22}	0.106384272	0.999775010	0.106408213	0.025982421	0.999770320	0.025988390	0.000000122	0.999761903	0.000000122
Γ_{11}	0.000000807	0.999139475	0.000000807	0.232206423	0.999308548	0.232367093	0.000014251	0.999308051	0.000014261
Γ_{12}	0.232206423	0.999308548	0.232367093	0.000014251	0.999308051	0.000014261	0.120432754	0.999308234	0.120516123
Γ_{13}	0.002614437	0.997571500	0.002620802	0.000317048	0.997580698	0.000317817	0.018275475	0.997622775	0.018319023
Γ_{23}	0.097842498	0.998931656	0.097947139	0.003639826	0.998934764	0.003643707	0.003657164	0.998926845	0.003661092
$T = \frac{A}{B}$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$ eq. (6.27)									
$\psi_0 = F_{1-\alpha}(m, n - m) = 2.604964777$ eq. (6.28) with $\alpha = 0.05$, $m = 3$, and $n = 400 \times 400$									

Table 7.4: CASE II: Hypothesis testing after 50 iterations in cases of the 2m x 2m x 1000m anomaly at the azimuthal orientations of 90°, 45°, and 0° in 0.01E²/Hz white noise.

Orient	SNR					
	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
90°	0.093	0.015	0.089	0.014	0.088	0.021
45°	0.092	0.048	0.048	0.044	0.065	0.065
0°	0.093	0.089	0.015	0.014	0.021	0.088
SNR = (the rms of gradient signal)/(the rms of noise background)						

Table 7.5: CASE II: The ratio of the gradient signals due to the 2m x 2m x 1000m anomaly at 50m depth and $0.01E^2/\text{Hz}$ white noise.

Orient.	Iteration	$\tilde{\Theta}_i - \Theta_0$	N_i	$C_{\Theta}^{-1}[C_{\Theta} + N_i^{-1}]C_{\Theta}^{-1}$	$T = \frac{\mathbf{R}/m}{\mathbf{\Omega}/n - m}$
90°	1	$\begin{bmatrix} 14.445 \\ -10.388 \\ -0.194 \end{bmatrix}$	$\begin{bmatrix} 8.301E-003 & 1.429E-006 & 0 \\ 1.429E-006 & 3.388E-003 & 0 \\ 0 & 0 & 1.169E-002 \end{bmatrix}$	$\begin{bmatrix} 1.259E-003 & -6.274E-008 & 0 \\ -6.274E-008 & 1.475E-003 & 0 \\ 0 & 0 & 86.542 \end{bmatrix}$	$\frac{1.228533782}{0.999319425}$
	2	$\begin{bmatrix} 3.168 \\ -3.744 \\ -1.902 \end{bmatrix}$	$\begin{bmatrix} 8.438E-003 & 1.700E-006 & 0 \\ 1.700E-006 & 3.428E-003 & 0 \\ 0 & 0 & 2.369E-2 \end{bmatrix}$	$\begin{bmatrix} 1.257E-003 & -7.254E-008 & 0 \\ -7.254E-008 & 1.471E-003 & 0 \\ 0 & 0 & 85.266 \end{bmatrix}$	$\frac{1.040186388}{0.999315835}$
45°	1	$\begin{bmatrix} 31.084 \\ -17.178 \\ -1.378 \end{bmatrix}$	$\begin{bmatrix} 7.202E-002 & -6.970E-002 & 0 \\ -6.970E-002 & 7.202E-002 & 0 \\ 0 & 0 & 1.440E-001 \end{bmatrix}$	$\begin{bmatrix} 1.381E-003 & 2.608E-004 & 0 \\ 2.608E-004 & 1.381E-003 & 0 \\ 0 & 0 & 7.942 \end{bmatrix}$	$\frac{5.515943597}{0.999398311}$
	2	$\begin{bmatrix} -7.232 \\ 4.164 \\ -1.238 \end{bmatrix}$	$\begin{bmatrix} 7.665E-002 & -7.420E-002 & 0 \\ -7.420E-002 & 7.665E-002 & 0 \\ 0 & 0 & 1.533E-001 \end{bmatrix}$	$\begin{bmatrix} 1.366E-003 & 2.475E-004 & 0 \\ 2.475E-004 & 1.367E-003 & 0 \\ 0 & 0 & 7.522 \end{bmatrix}$	$\frac{3.876353895}{0.999359939}$

Table 7.6: CASE II: The example of the test statistic operations for Γ_{12} with respect to the azimuthal orientations of 90° and 45°.

7.4 CASE III: Predicted location and fixed orientation using individual gradient components in $1E^2/\text{Hz}$ (or $300E^2/(\text{cyc}/\text{m})^2$) white noise environment

This example investigates the capability of the parameter estimation when the noise background increases. Consider the gradient signals due to the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly in $1E^2/\text{Hz}$ (or $300E^2/(\text{cyc}/\text{m})^2$) white noise environment. Here, the flat psd is $300E^2/(\text{cyc}/\text{m})^2$ for all frequencies. The initial value of Θ_0 is set to zero. The a priori covariance matrix of Θ is taken from (7.4). The depth variances $\sigma_D^2 = 1\text{m}^2$ and $\sigma_D^2 = 100\text{m}^2$ are considered for the estimation.

We begin with the choice of $\sigma_D^2 = 1\text{m}^2$. Figures 7.4 – 7.6 plot the horizontal location and depth errors as well as the corresponding test statistics associated with $\beta' = 90^\circ, 45^\circ, \text{ and } 0^\circ$, respectively. For $\beta' = 90^\circ$, the east coordinate errors for all components vary from 1m to 560m through the iterative procedure. The component Γ_{12} yields the largest error of about 560m at iteration 50. The north coordinate is accurately estimated within 10m using $\Gamma_{33}, \Gamma_{11}, \text{ and } \Gamma_{13}$ after 15 iterations. For the depth estimation, the components Γ_{33} and Γ_{11} yield the errors that are about 5m smaller than using the other components.

For $\beta' = 45^\circ$, the values of the north errors are about 15-70m after 50 iterations, shown in Figure 7.5. The east and north errors vary in the range of 10-300m and oscillate towards convergence (except in the Γ_{22} case). At the end of this section, we shall show that the choice of a priori variance $\sigma_D^2 = 100\text{m}^2$ provides more stable results.

Examples of $\text{MSPE}\{\tilde{\Theta}_{50}\}$ at iteration 50 for Γ_{33} obtained from eq. (7.2) are summarized as follows:

I: The azimuthal orientation is 90° and $\hat{\sigma}_0^2 = 1.004195$

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 1.004195 \begin{bmatrix} 98.252285 & -0.000347 & 0.000000 \\ -0.000347 & 726.542725 & 0.000000 \\ 0.000000 & 0.000000 & 0.990754 \end{bmatrix} [\text{m}]^2 \quad (7.8)$$

II: The azimuthal orientation is 45° and $\hat{\sigma}_0^2 = 1.004178$

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 1.004178 \begin{bmatrix} 419.723839 & 315.857464 & 0.000000 \\ 315.857464 & 419.664694 & 0.000000 \\ 0.000000 & 0.000000 & 0.991308 \end{bmatrix} [\text{m}]^2 \quad (7.9)$$

III: The azimuthal orientation is 0° and $\hat{\sigma}_0^2 = 1.004202$

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 1.004202 \begin{bmatrix} 59.206306 & -0.000009 & 0.000000 \\ -0.000009 & 1.827906 & 0.000000 \\ 0.000000 & 0.000000 & 1.749041 \end{bmatrix} [\text{m}]^2 \quad (7.10)$$

The MSPE's indicate that the north and east estimates, especially where the anomaly is at $\beta' = 45^\circ$, are mutually correlated and independent of the depth. Finally, the rmse's of the adjusted estimates obtained from the diagonal elements of eq. (7.2) are summarized in Table 7.7.

As stated in the previous case, since the signals are small compared to the strong noise, Figures 7.4b – 7.6b show that the values of test statistic do not provide useful information in the hypothesis testing after 50 iterations.

	rmse (m)					
	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
The azimuthal orientation fixed to 90°						
Northing	9.933	29.694	3.984	29.171	9.394	14.873
Easting	27.011	29.097	24.840	29.640	28.065	13.849
Depth	0.997	1.000	0.971	1.000	0.994	0.995
The azimuthal orientation fixed to 45°						
Northing	20.530	25.997	17.166	19.208	20.757	19.914
Easting	20.528	24.923	18.317	19.211	21.584	19.160
Depth	0.998	0.997	0.989	0.989	0.995	0.992
The azimuthal orientation fixed to 0°						
Northing	27.003	28.631	28.981	27.727	26.743	28.083
Easting	9.990	9.405	29.661	24.355	26.748	9.235
Depth	0.997	0.995	1.001	0.999	0.999	0.994

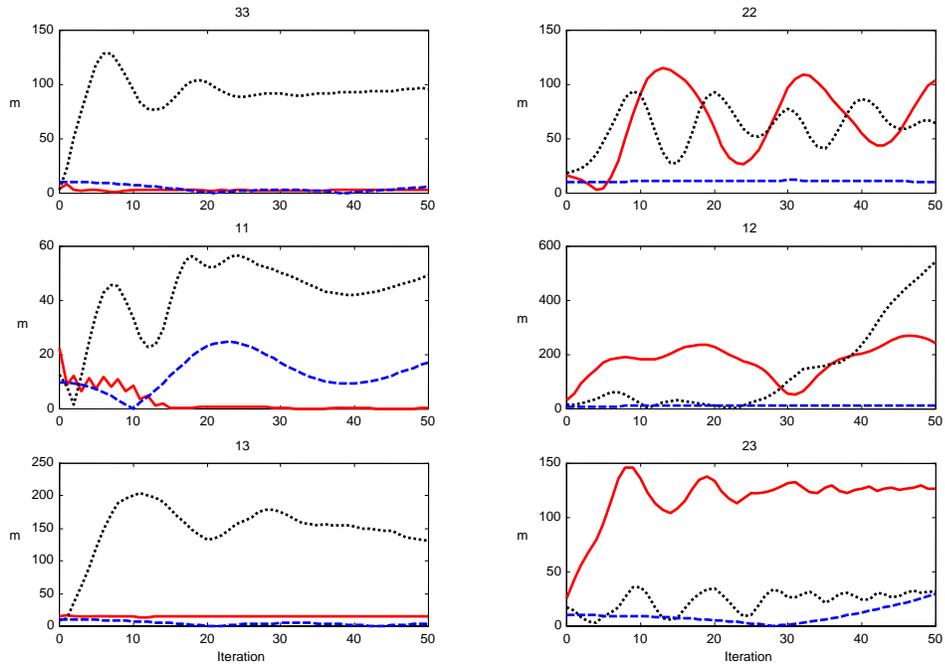
Table 7.7: CASE III: The root-mean-square-error (rmse) of the adjusted parameters after 50 iterations

The values of the test statistic after iterations 50 are summarized in Table 7.8. These values for all gradient components are below the threshold ψ_0 for $\alpha = 0.05$; the null hypothesis, H_0 , is accepted. Thus, $\tilde{\Theta}$ and Θ_0 are statistically equal according to (6.24). Since the gradient signals are embedded in strong noise (i.e., $1E^2/\text{Hz}$ white noise), the test statistic is not able to identify the inconsistency between $\tilde{\Theta}$ and Θ_0 unless a complete reduction of noise backgrounds has been done. Table 7.9 shows, for instance, the SNR for Γ_{33} at $\beta' = 90^\circ$ (whose signal magnitude is largest) is only $9.243E-3$ (even smaller, for Γ_{12} ; it is 15% of Γ_{33}). The numerical analyses of CASE II provide a good comparison.

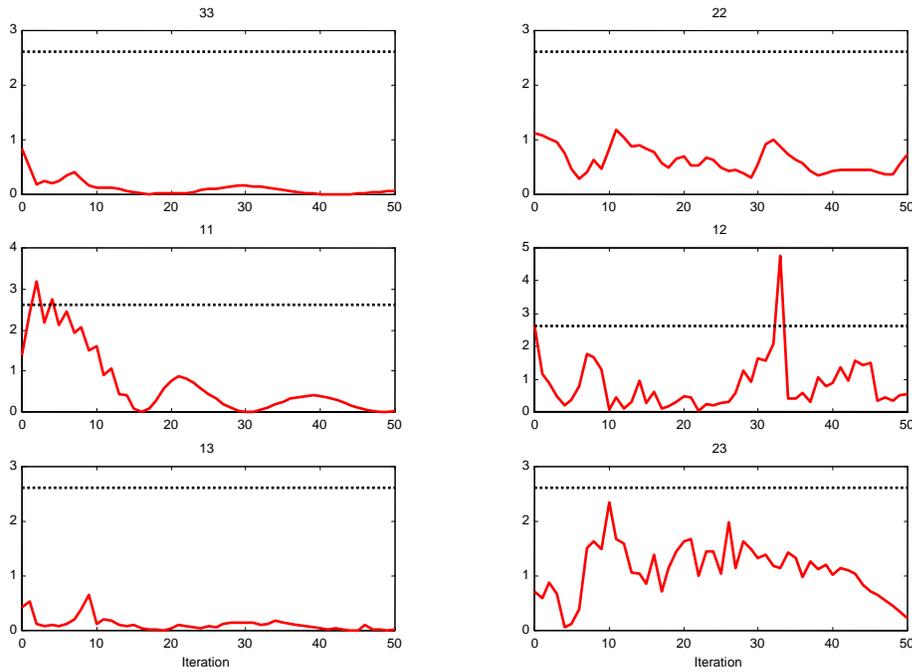
One can compare the results discussed above with those in CASE II where the white noise level is only $0.01E^2/\text{Hz}$.

The adjusted estimates stabilize with respect to the iterative procedure, $|\tilde{\Theta}_i - \tilde{\Theta}_{i-1}| \rightarrow 0$ with the updated $\Theta_0 = \tilde{\Theta}_{i-1}$ (as the a priori information for the next iteration, i.e., the i^{th} iteration), and in turn yield the relative increase \mathbf{R} of (6.16) and subsequent smaller test statistics.

Referring to the choice of $\sigma_D^2 = 1\text{m}^2$, we have seen that the depth errors oscillate with the iterations (Figures 7.4a-7.6a). These results occur due to overweighting of the depth parameter whereas its initial approximation is assigned to 60m (the true depth = 50m). In comparison, Figure 7.7 provides an example of the estimation for the case of $\beta' = 90^\circ$ with the choice of $\sigma_D^2 = 100\text{m}^2$. The plots show that selecting $\sigma_D^2 = 100\text{m}^2$ affects the adjusted solutions. However, the north error using Γ_{12} continually grows with respect to the number of iterations. The next sections show how the adjusted solutions can be improved using all six combined gradient components.

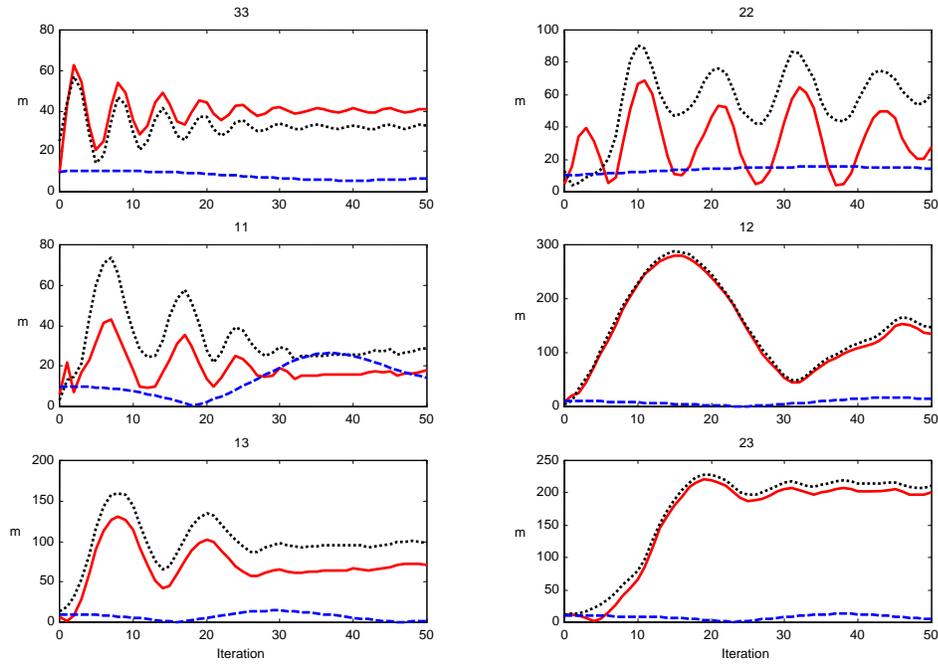


(a)

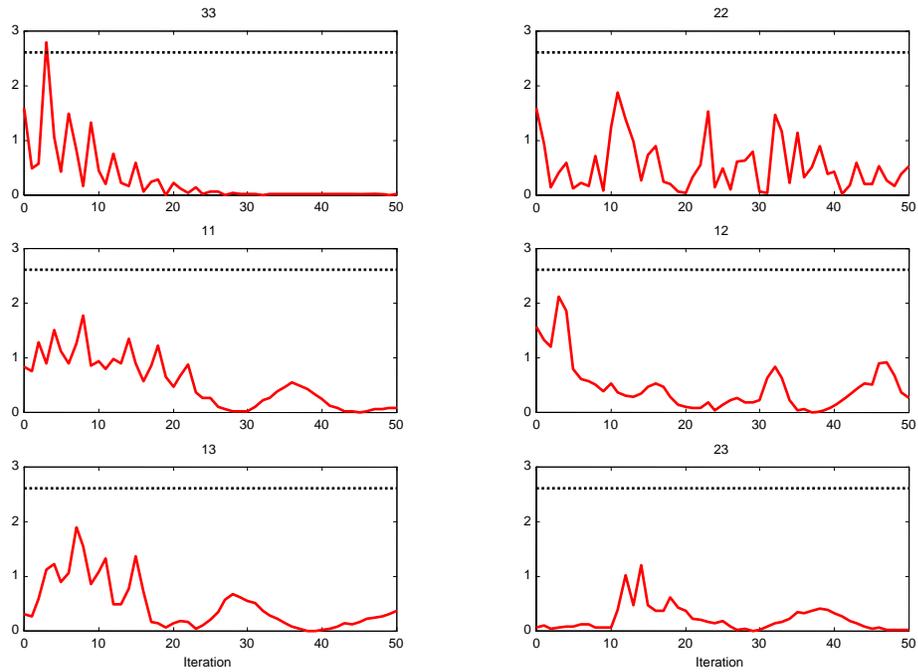


(b)

Figure 7.4: CASE III: (a) The absolute errors of north (red solid line), east (black dotted line), and depth (blue dashed line) at the azimuthal orientation of 90° ; (b) Test statistic (red solid line) and $\psi_0 = 2.60$ (black dotted line) with $m = 3$ and $n = 400 \times 400$.

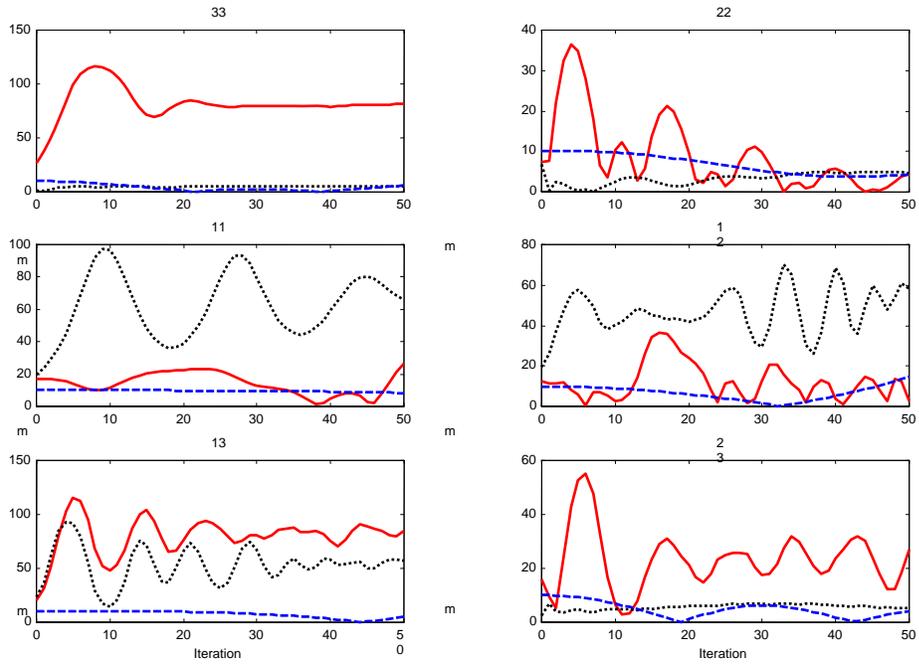


(a)

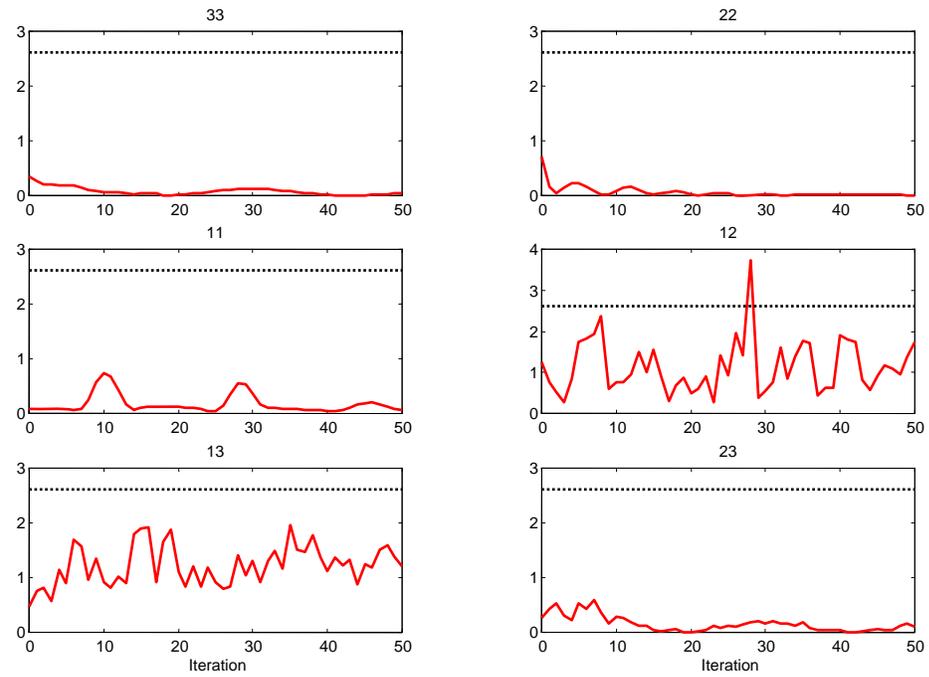


(b)

Figure 7.5: CASE III: (a) The absolute errors of north (red solid line), east (black dotted line), and depth (blue dashed line) at the azimuthal orientation of 45° ; (b) Test statistic (red solid line) and $\psi_0 = 2.60$ (black dotted line) with $m = 3$ and $n = 400 \times 400$.



(a)



(b)

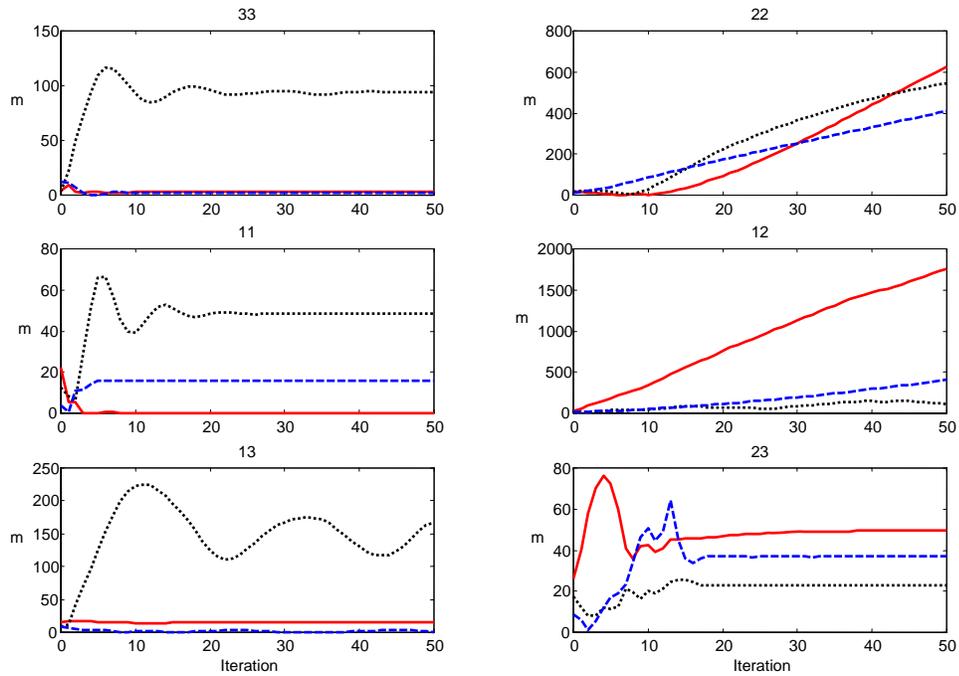
Figure 7.6: CASE III: (a) The absolute errors of north (red solid line), east (black dotted line), and depth (blue dashed line) at the azimuthal orientation of 0° ; (b) Test statistic (red solid line) and $\psi_0 = 2.60$ (black dotted line) with $m = 3$ and $n = 400 \times 400$.

Grad.	Test statistic								
	90°			45°			0°		
	A	B	T	A	B	T	A	B	T
Γ_{33}	0.064350953	1.004212279	0.064081026	0.010591046	1.004196414	0.010546787	0.047610305	1.004220127	0.047410228
Γ_{22}	0.738065812	1.000573658	0.737642657	0.540727323	1.000562533	0.540423317	0.010036160	1.000591197	0.010030230
Γ_{11}	0.031714186	1.001325206	0.031672214	0.089353678	1.001338422	0.089234245	0.058996333	1.001360611	0.058916171
Γ_{12}	0.536019127	1.000548146	0.535725471	0.260545184	1.000566849	0.260397578	1.733903010	1.000537582	1.732971395
Γ_{13}	0.027702024	0.998572817	0.027741616	0.366835329	0.998555888	0.367365847	1.192317190	0.998559085	1.194037696
Γ_{23}	0.220645007	0.998676353	0.220937450	0.028394839	0.998681319	0.028432332	0.100707828	0.998712830	0.100837623
$T = \frac{A}{B}$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$ eq. (6.27)									
$\psi_0 = F_{1-\alpha}(m, n - m) = 2.604964777$ eq. (6.28)									
with $\alpha = 0.05$, $m = 3$, and $n = 400 \times 400$									

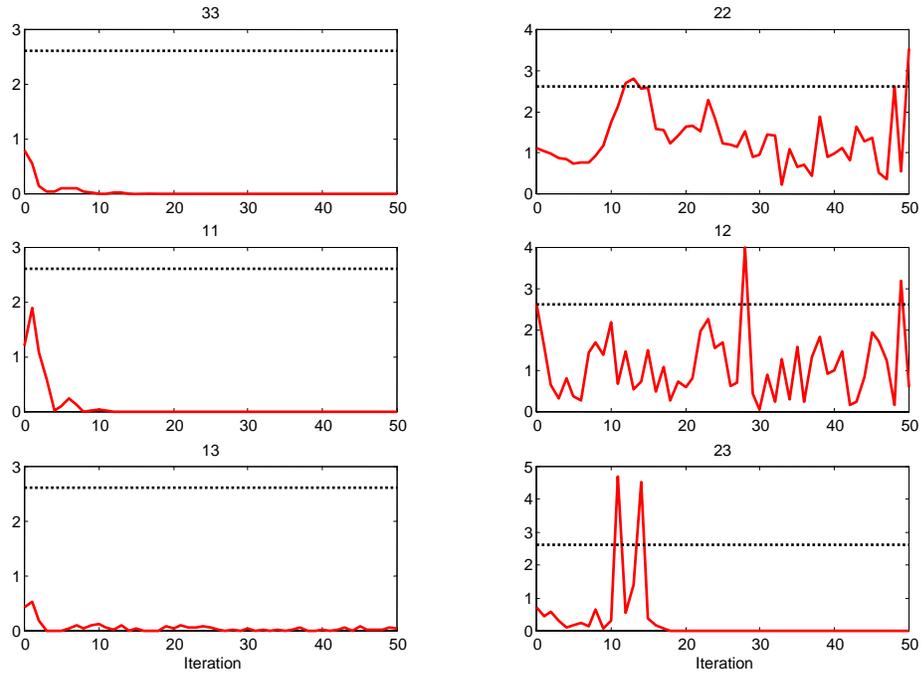
Table 7.8: CASE III: Hypothesis testing after 50 iterations in cases of the 2m x 2m x 1000m anomaly at the azimuthal orientations of 90°, 45°, and 0° in 1E²/Hz white noise environment.

Orient	SNR					
	Γ_{33}	Γ_{22}	Γ_{11}	Γ_{12}	Γ_{13}	Γ_{23}
90°	9.243E-3	1.461E-3	8.909E-3	1.439E-3	8.840E-3	2.052E-3
45°	9.146E-3	4.802E-3	4.800E-3	4.350E-3	6.481E-3	6.480E-3
0°	9.243E-3	8.913E-3	1.461E-3	1.439E-3	2.052E-3	8.828E-3
SNR = (the rms of the gradient signal)/(the rms of the noise background)						

Table 7.9: CASE III: The ratio of the gradient signals due to the 2m x 2m x 1000m anomaly at 50m depth and 1E²/Hz white noise.



(a)



(b)

Figure 7.7: CASE III: (a) The absolute errors of north (red solid line), east (black dotted line), and depth (blue dashed line) at the azimuthal orientation of 90° with $\sigma_D^2 = 100\text{m}^2$; (b) Test statistic (solid line) and $\psi_0 = 2.60$ (dotted line) with $m = 3$ and $n = 400 \times 400$.

7.5 CASE IV: Predicted location and fixed orientation using the combination of all six gradient components in $0.01E^2/\text{Hz}$ (or $3E^2/(\text{cyc}/\text{m})^2$) white noise environment

Consider only the case where the azimuthal orientation, 90° , of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly is assumed known perfectly (i.e., no orientation estimation). We test the robustness of the random effects model using all six combined gradients in $0.01E^2/\text{Hz}$ white noise environment. The initial value of Θ_0 is zero and we choose $\sigma_D^2 = 1\text{m}^2$. The parameter estimation pertains to the case of multiple data tracks and multiple sensors (MM) in Section 6.3.4. We consider no correlation between sensors in the white noise environment.

Figure 7.8 shows the numerical results for all combined gradients versus the results of using individual components obtained from CASE II. Overall, the combination provides an improvement in the estimation. The north errors decrease to 0.17m after 50 iterations. Although the combination does not give the smallest error in the east coordinate, about 0.5m , it is superior to the estimation using the other components except the component Γ_{11} for which the error is about 0.11m . The depth errors become stabilized after 5 iterations; these results imply that the iterated solutions, $\tilde{\Theta}_i$, converge faster than those from using the individual gradients.

The MPSE of $\tilde{\Theta}_{50}$, obtained from (7.2) and $\hat{\sigma}_0^2 = 0.998846$, is:

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 0.998846 \begin{bmatrix} 0.216992 & -0.000017 & 0 \\ -0.000017 & 8.354125 & 0 \\ 0 & 0 & 0.174644 \end{bmatrix} [\text{m}]^2 \quad (7.11)$$

rmse(m)		
Northing	Easting	Depth
0.465	2.889	0.418

Table 7.10: CASE IV: The root-mean-square-error (rmse) of the adjusted parameters after 50 iterations

Taken from (7.11), the rmse of the east coordinate in Table 7.10 significantly decreases, compared to the rmse using the individual gradients in Table 7.3.

Table 7.11 describes the acceptance of the null hypothesis statistically indicating the equivalency of $\tilde{\Theta}$ and Θ_0 after 50 iterations at $\alpha = 0.05$ level of significance. In Figure 7.9, after only four iterations, the test statistic is below the threshold. Therefore, the numerical analyses show that the use of the six gradient combination leads to an improvement of parameter estimation by the random effects model.

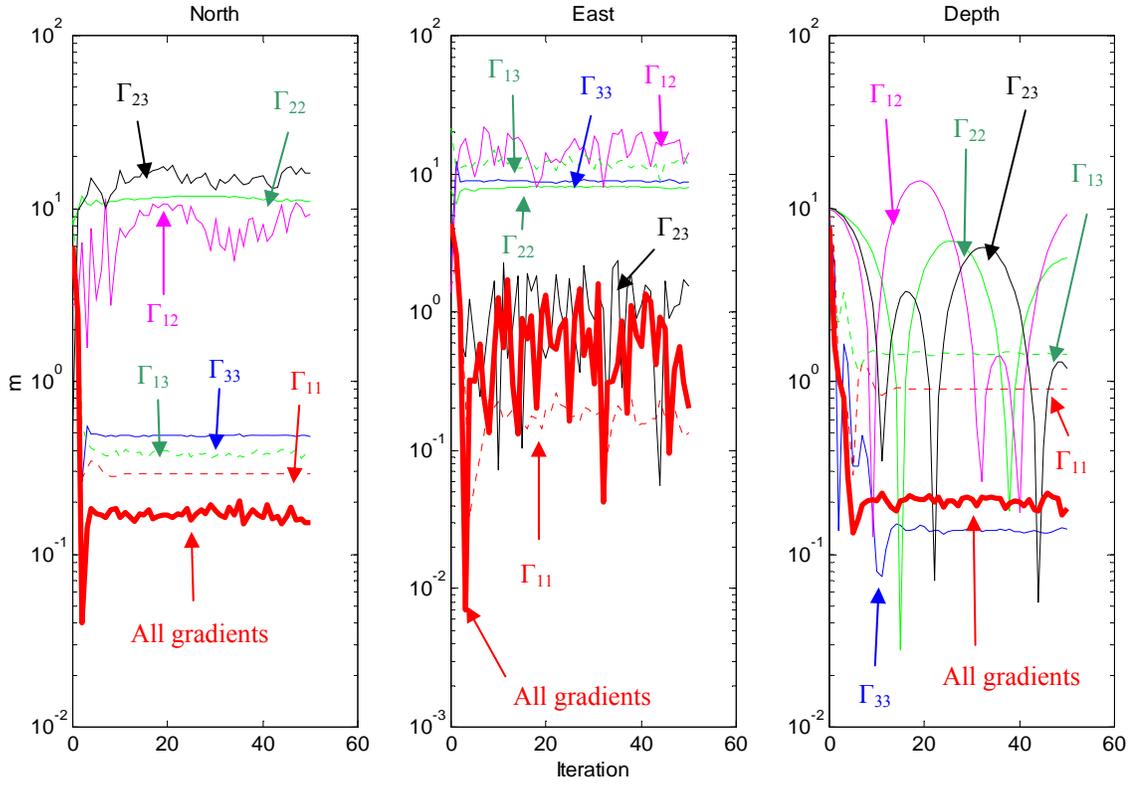


Figure 7.8: CASE IV: The absolute errors at the azimuthal orientation of 90° in $0.01E^2/\text{Hz}$ white noise.

	Test statistic		
	A	B	T
All gradients	0.001219191	0.998849044	0.001220598
$T = \frac{A}{B}$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$			eq. (6.27)
$\psi_0 = F_{1-\alpha}(m, n - m) = 2.604918546$			eq. (6.28)
with $\alpha = 0.05$ $m = 3$ and $n = 6 \times 400 \times 400$			

Table 7.11: CASE IV: Hypothesis testing after 50 iterations for the case of the $2m \times 2m \times 1000m$ anomaly at the azimuthal orientation of 90° in $0.01E^2/\text{Hz}$ white noise

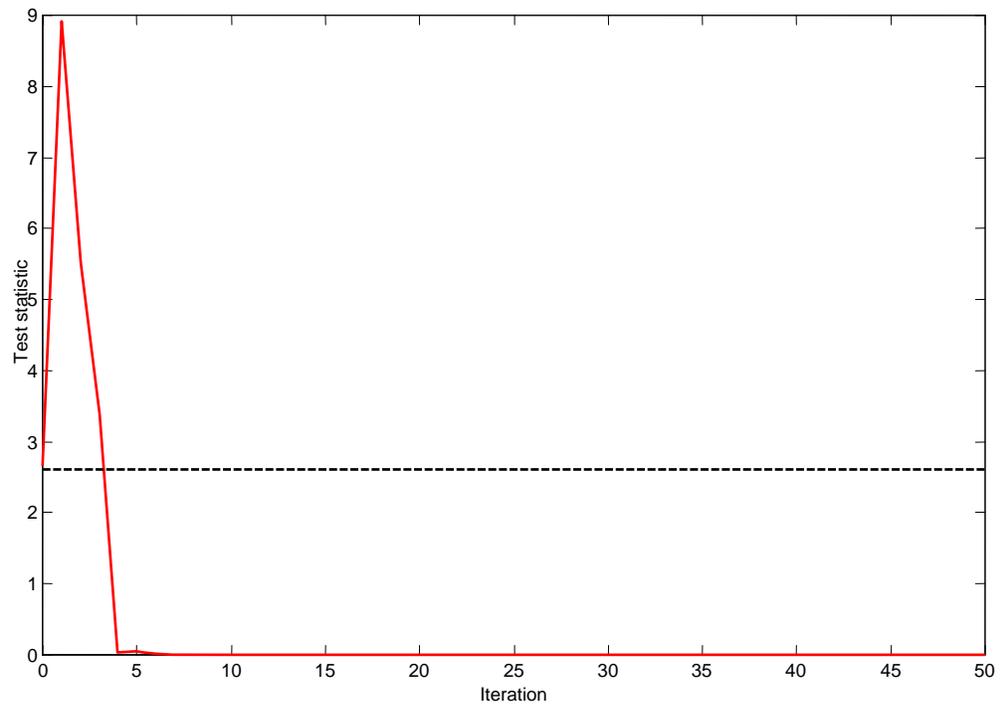


Figure 7.9: CASE IV: The test statistic (red solid line) and threshold $\psi_0 = 2.60$ (black dashed line) for a given $\alpha = 0.050$.

7.6 CASE V: Predicted location and fixed orientation using the combination of all six gradient components in $1E^2/Hz$ (or $300E^2/(cyc/m)^2$) white noise environment

Suppose the output of matched filtering for $\beta' = 90^\circ$ is given as in CASE IV together with an a priori covariance matrix corresponding to $1E^2/Hz$ white noise. We consider no correlation between sensors in the white noise environment. Figure 7.10 shows the results of all six combined gradients versus the adjusted results of using individual components obtained from CASE III. Note that although the combination does not provide the best estimates among the other components, its solutions are much more stable for the north and east estimates. The estimates of depth, though, either using the combination or individual components, yield small fluctuation during the iterative process. It means that the choice of the a priori variance $\sigma_D^2 = 1m^2$ may not be applicable when the noise is high in association with the initial approximation of depth = 60m. Furthermore, a larger number of iterations may be needed due to a slow convergence of the solution.

We thus use $\sigma_D^2 = 100m^2$ instead to relax the constraint of (6.4). The results are drawn in Figure 7.11. The plot also includes the results of individual gradients where $\sigma_D^2 = 100m^2$ is employed, taken from Figure 7.7. Note that the depth estimation is more stable, as compared to the case $\sigma_D^2 = 1m^2$ in Figure 7.10. The rapid stability of the solution renders the values of the test statistic significantly smaller after the first 2 iterations which then quickly approach zero; whereas, for $\sigma_D^2 = 1m^2$, the test statistic oscillates and only gradually diminishes with the iterations, see Figures 7.10b versus 7.11b. (Note that the depth solution in Figure 7.10a causes the oscillation of the test statistic in Figure 7.10b.) The rmse's of the adjusted parameters, obtained from (7.2), and the test statistic computation are summarized in Tables 7.12 and 7.13, respectively.

Choice	rmse(m)			$\hat{\sigma}_0^2$
	Northing	Easting	Depth	
$\sigma_D^2 = 1m^2$	4.106	19.821	0.972	1.000656
$\sigma_D^2 = 100m^2$	4.270	20.181	3.919	1.000656

Table 7.12: CASE V: The root-mean-square-error (rmse) of the adjusted parameters using all six combined gradients and the a posteriori variances after 50 iterations.

Choice	Test statistic		
	A	B	T
$\sigma_D^2 = 1\text{m}^2$	0.000254465	1.000659663	0.000254297
$\sigma_D^2 = 100\text{m}^2$	0.004578409	1.000659439	0.004575392
$T = \frac{A}{B}$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$			eq. (6.27)
$\Psi_0 = F_{1-\alpha}(m, n - m) = 2.604918546$			eq. (6.28)
with $\alpha = 0.05$ $m = 3$ and $n = 6 \times 400 \times 400$			

Table 7.13: CASE V: Hypothesis testing after 50 iterations for the case of the 2m x 2m x 1000m anomaly at the azimuthal orientation of 90° in 1E²/Hz white noise using all six combined gradients.

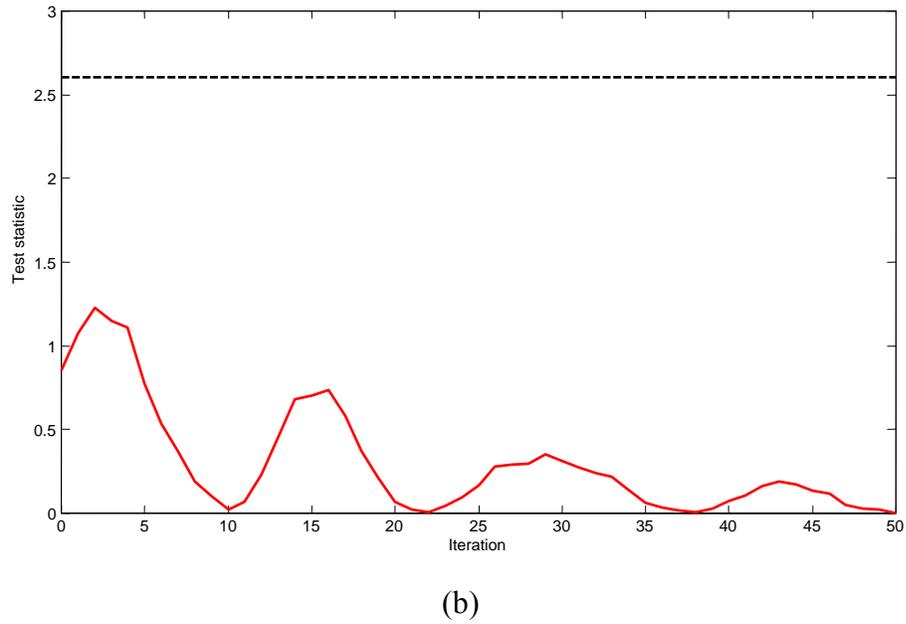
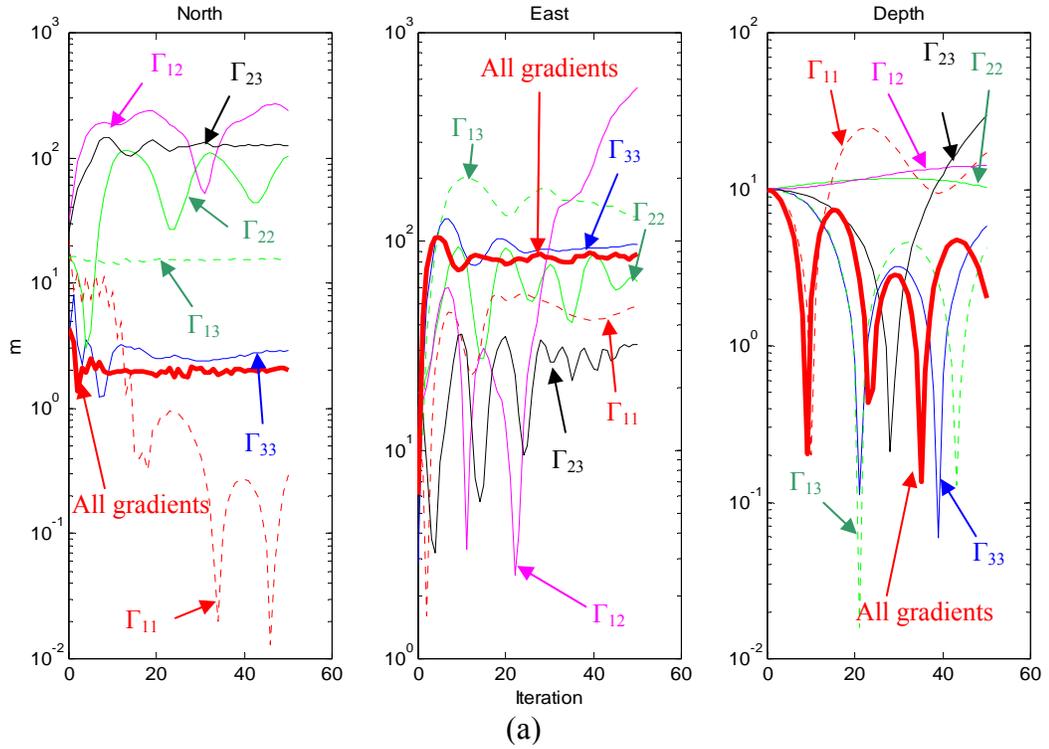


Figure 7.10: CASE V: (a) The absolute errors at the azimuthal orientation of 90° using all six combined gradients in $1E^2/\text{Hz}$ white noise with $\sigma_D^2 = 1\text{m}^2$; (b) The test statistic (red solid line) and threshold $\psi_0 = 2.60$ (black dashed line) for a given $\alpha = 0.050$.

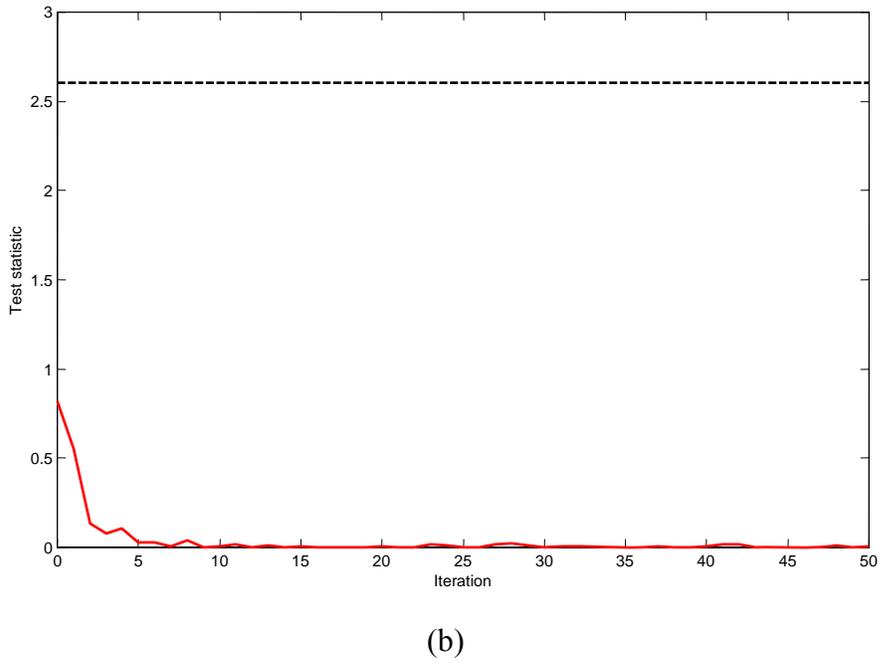
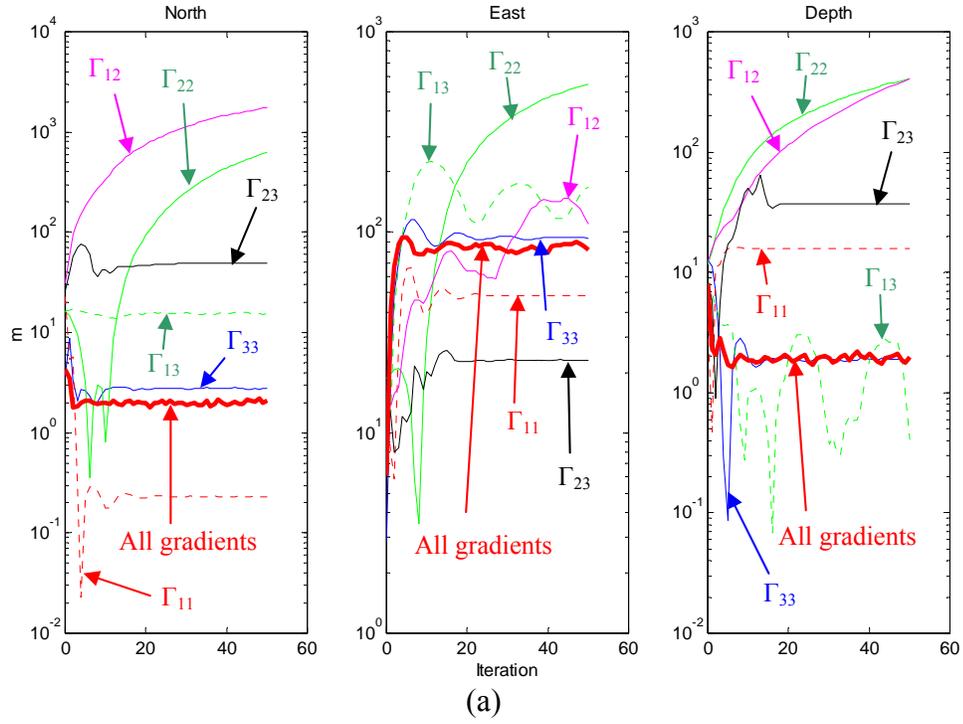


Figure 7.11: CASE V: (a) The absolute errors at the azimuthal orientation of 90° using all six combined gradients in $1E^2/\text{Hz}$ white noise with $\sigma_D^2 = 100\text{m}^2$; (b) The test statistic (red solid line) and threshold $\psi_0 = 2.60$ (black dash line) for a given $\alpha = 0.050$.

7.7 CASE VI: Predicted location and fixed orientation using individual gradients in geological background

For this case, we add the full geological background plus $1E^2/\text{Hz}$ white noise to the gradient signals due to the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly at $\beta' = 90^\circ$. We use the same a priori information as in the previous cases except that only the a priori variance of depth $\sigma_D^2 = 100\text{m}^2$ is used. The parameter estimation is conducted independently by using different gradient components. The physical correlations between observations are taken into account and eq. (A.1) in Appendix A is applied (the instrumental noise N_F in Table A.1 is also included). Figures 7.12 through 7.14 show the iterated solutions of north and east coordinates and depth, respectively, using the different psd models shown in Appendix A. We compare the iterative solutions before and after the subtraction of a geological model described in Chapter 5 (called WB and WOB respectively) using Psd models 1 (black dotted line) and 2 (red solid line). We also show that the small change of Psd model 1 (i.e., we slightly change the parameters of Psd model 1) affects the parameter estimation.

Overall, the numerical results of Γ_{33} , Γ_{11} , Γ_{13} , and Γ_{23} for WB and WOB stabilize except the east coordinate error of Γ_{13} after 15 iterations. For other components, the north, east, and depth errors steadily increase with respect to the number of iterations.

The numerical results shown in Figures 7.12 – 7.14 indicate that the parameter estimation using one gradient component, for the WOB case using Psd model 2, does not give a better result of estimating coordinates than that for the WB using Psd model 1. For instance, after 50 iterations in Figure 7.12, the north error with Γ_{13} and the WB case is 42.501m (black-dotted line) whereas the error is 199.460m for WOB (red-solid line). However, using Γ_{33} in the WB case, the north error is 4.750m; while using Γ_{33} in the WOB case, the error is only 1.061m. One possible reason is that the psd's may not be accurately modeled. To support this statement, we slightly change the parameters of Psd model 1 to see how it affects the results by replacing $\sigma_{10}^2 = 2.0 \times 10^{-10} \text{ m}^4/\text{s}^4$ and $\sigma_{11}^2 = 3.0 \times 10^{-13} \text{ m}^4/\text{s}^4$. After 50 iterations, the numerical results show that, for instance, the north error by using Γ_{13} largely increases to several thousand meters (see blue dashed line in Figure 7.12).

The Γ_{22} , Γ_{12} , and Γ_{23} signals are small with respect to the 90° orientation of the anomaly and can be entirely corrupted by noise. For example, the rms of the signal Γ_{12} ($0.831E-3$ Eötvös, taken from Table 7.1) is only about 0.02% of the full geological background plus $1E^2/\text{Hz}$ white noise (4.005 Eötvös). As a consequence, the location errors of Γ_{22} , Γ_{12} , and Γ_{23} are large, about 10-100m.

Examples of $\text{MSPE}\{\tilde{\Theta}_{50}\}$ for Γ_{33} at iteration 50 obtained from (7.2) are:

(a) The geological model was not removed (and Psd model 1 was used) and $\hat{\sigma}_0^2 = 16.843036$:

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 16.843036 \begin{bmatrix} 75.745876 & 0.000052 & 0 \\ 0.000052 & 698.254406 & 0 \\ 0 & 0 & 44.619035 \end{bmatrix} [\text{m}^2] \quad (7.12)$$

(b) The geological model was not removed (and Psd model 1 was changed) and $\hat{\sigma}_0^2 = 17.901517$:

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 17.901517 \begin{bmatrix} 86.585646 & -0.001380 & 0 \\ -0.001380 & 714.178836 & 0 \\ 0 & 0 & 48.245677 \end{bmatrix} [\text{m}^2] \quad (7.13)$$

(c) The geological background was removed (and Psd model 2 was used) and $\hat{\sigma}_0^2 = 13.087721$:

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 13.087721 \begin{bmatrix} 77.354893 & -0.000513 & 0 \\ -0.000513 & 693.579463 & 0 \\ 0 & 0 & 45.152834 \end{bmatrix} [\text{m}^2] \quad (7.14)$$

We assess the success of the parameter estimation. At iteration 50, the values of the test statistic in Table 7.14 are below the threshold ψ_0 at $\alpha = 0.05$ level of significance; the null hypotheses are accepted. Although the estimated locations do not coincide with the true locations, especially the results of using Γ_{22} , Γ_{12} , and Γ_{23} , they equal Θ_0 statically. As stated in the previous cases, since the anomaly signals are small, compared to the noise background, the test statistic is not able to identify for large location errors. It should be noted also that the values of $\mathbf{\Omega}/(n - m)$ are large, compared to \mathbf{R}/m , especially, for component Γ_{22} . These results are due to an incomplete reduction of noise backgrounds.

Of interest is, assuming the a priori variance component to be unity, i.e., $\sigma_0^2 = 1$, whether or not the a posteriori variance component without constraint, i.e., $\bar{\sigma}_0^2 = \mathbf{\Omega}/n - m$, in the adjustment is significantly different from its a priori value.

With the χ^2 -distribution, where $\frac{(n - m)\bar{\sigma}_0^2}{\sigma_0^2} \sim \chi_{n-m}^2$ holds, the two-tailed hypothesis test can be set up as follows (Leick, 1997, pp. 142-143, and Kock, 1999, pp. 286-287):

$$H_0: \sigma_0^2 = \bar{\sigma}_0^2 \quad \text{versus} \quad H_1: \sigma_0^2 \neq \bar{\sigma}_0^2 \quad (7.15)$$

It will be rejected with a significance level of α if the test statistic, T_χ ,

$$T_\chi > \chi^2_{1-\frac{\alpha}{2}, n-m} \text{ or } T_\chi < \chi^2_{\frac{\alpha}{2}, n-m} \text{ with } T_\chi = \frac{(n-m)\bar{\sigma}_0^2}{\sigma_0^2} = \frac{\mathbf{\Omega}}{\sigma_0^2} \quad (7.16)$$

For our case, the rejection of H_0 is taken as an indicator that the reduction of the noise background in the observation \mathbf{y} in eq. (6.13) is not significant. The results of the χ^2 test are given in Table 7.14. For $\alpha = 0.05$, the hypothesis H_0 has been rejected for all gradients in either WB (using Psd model 1) or WOB (using Psd model 2) case. For instance, for Γ_{33} in WOB, $T_\chi = 26.949 \times 10^5 > \chi^2_{0.975} = 1.611 \times 10^5$. The specific test indicates that the full geological background plus $1E^2/\text{Hz}$ white noise needs to be reduced as much as possible so as to be accept the null hypothesis. To support this statement, we estimate the same parameters due to the same anomaly but in $0.01E^2/\text{Hz}$ white noise (as an ideal case). We only show the χ^2 test (see CASE II for the F-test). As shown in Table 7.14, the test statistics T_χ for all components fall inside the acceptance interval. (More details using χ^2 test for various noise cases are discussed in CASE IX.)

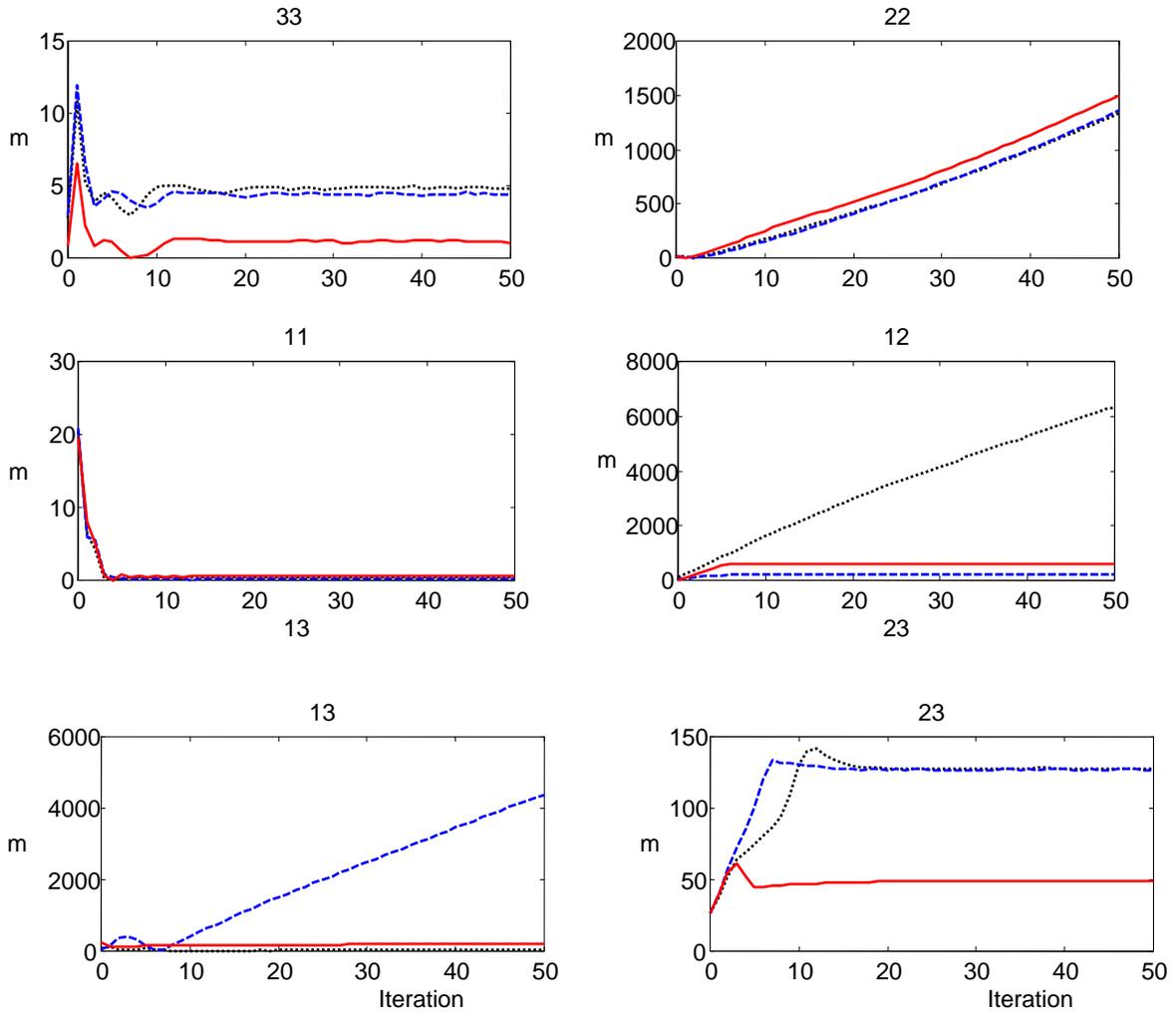


Figure 7.12: CASE VI: The absolute errors of north coordinates using individual gradient components in geological noise; black-dotted line indicates a geological model was not removed and Psd model 1 was used; blue-dashed line indicates a geological model was not removed and Psd model 1 was used but its parameters were slightly changed (see text); red-solid line indicates a geological model was removed and Psd model 2 was used.

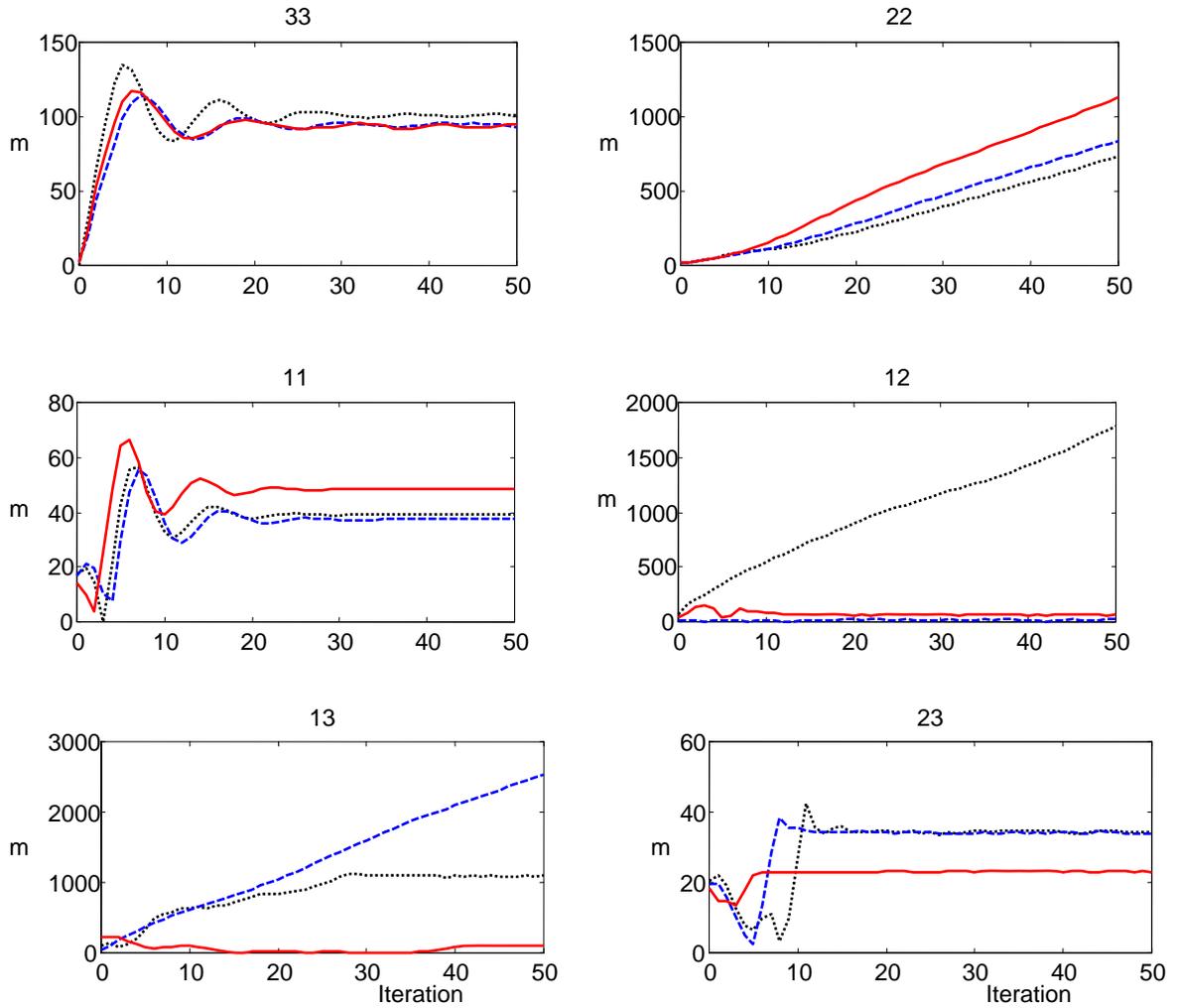


Figure 7.13: CASE VI: The absolute errors of east coordinate using individual gradient components in geological noise; black-dotted line indicates a geological model was not removed and Psd model 1 was used; blue-dashed line indicates a geological model was not removed and Psd model 1 was used but its parameters were slightly changed (see text); red-solid line indicates a geological model was removed and Psd model 2 was used.

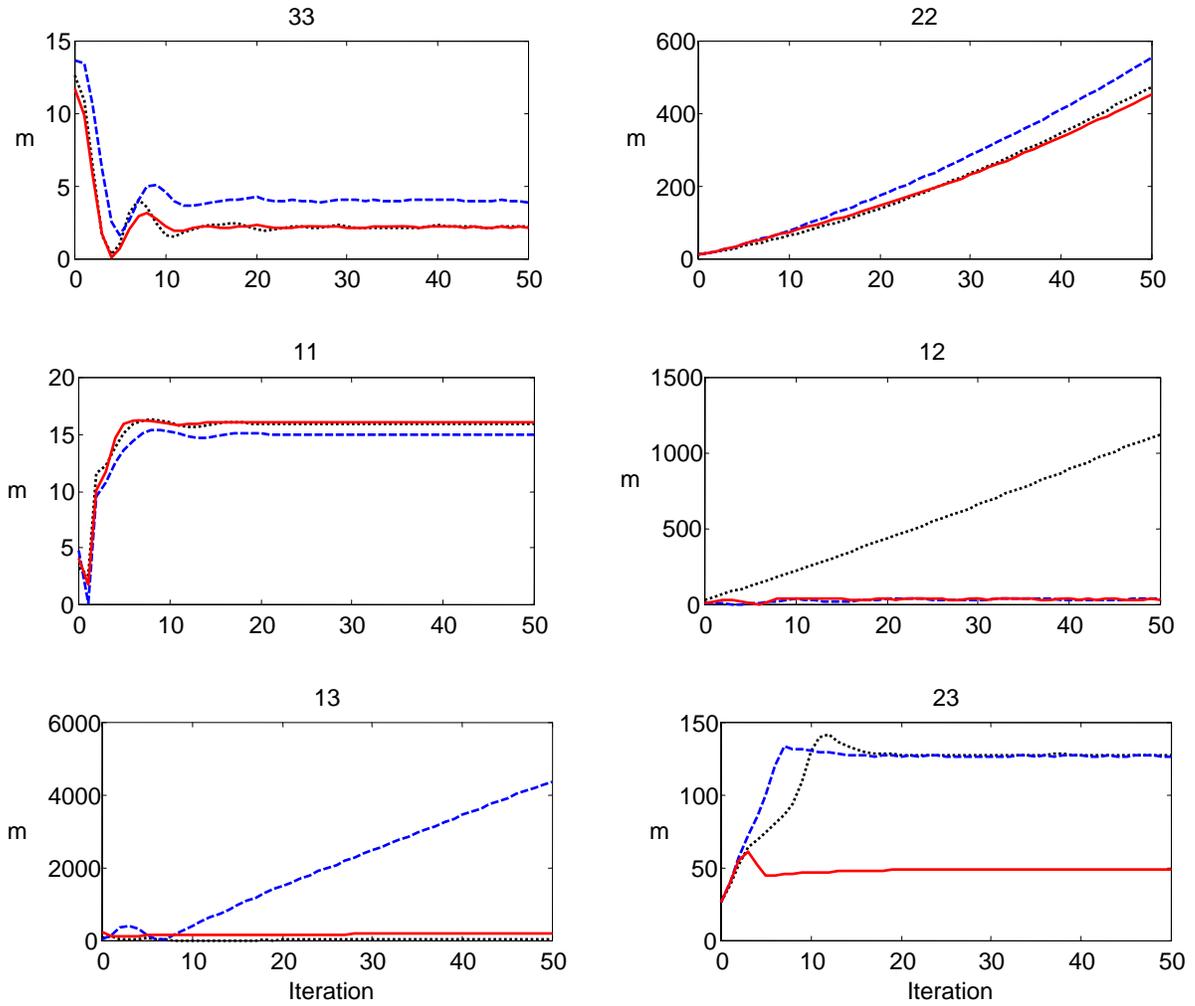


Figure 7.14: CASE VI: The absolute errors of depth using individual gradient components in geological noise; black-dotted line indicates a geological model was not removed and Psd model 1 was used; blue-dashed line indicates a geological model was not removed and Psd model 1 was used but its parameters were slightly changed (see text); red-solid line indicates a geological model was removed and Psd model 2 was used.

Grad	Test statistic									
	WB				WOB				WN	
	A	B	T	$T_\chi \times 10^{-5}$	A	B	T	$T_\chi \times 10^{-5}$	$T_\chi \times 10^{-5}$	
Γ_{33}	0.000005106	16.843352731	0.000000303	26.949	0.000027626	13.087966819	0.000002110	20.940	1.597	
Γ_{22}	12.329603691	90.663134823	0.135993573	145.058	17.249267597	84.536278081	0.204045742	135.256	1.599	
Γ_{11}	0.000000494	1.073199580	0.000000460	1.717	0.000009433	1.201789677	0.000007849	1.922	1.598	
Γ_{12}	3.279046755	10.836256313	0.302599593	17.338	0.088554138	10.869512897	0.008147019	17.390	1.600	
Γ_{13}	2.471165296	14.977656919	0.164990112	23.964	0.001416550	14.927722893	0.000094893	23.884	1.597	
Γ_{23}	0.088554138	17.696517774	0.000001760	28.313	0.000088649	17.263880658	0.000005134	27.622	1.597	

$T = A/B$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$ eq. (6.27)

$\psi_0 = F_{1-\alpha}(m, n - m) = 2.604964777$ eq. (6.28)

Reject H_0 if $T_\chi > \chi^2_{1-\frac{\alpha}{2}, n-m} = 1.611E+5$ or $T_\chi < \chi^2_{\frac{\alpha}{2}, n-m} = 1.589E+5$

with $T_\chi = (n - m)\bar{\sigma}_0^2/\sigma_0^2 = \mathbf{\Omega}/\sigma_0^2$, $\alpha = 0.05$, $m = 3$, $n = 400 \times 400$, and $\sigma_0^2 = 1$

WB = full geological background plus $1E^2/\text{Hz}$ white noise

WOB = full geological background plus $1E^2/\text{Hz}$ white noise minus a geological model

WN = $0.01E^2/\text{Hz}$ white noise

Table 7.14: CASE VI: Hypothesis testing via F test and χ^2 test after 50 iterations

7.8 CASE VII: Predicted location and fixed orientation using six combination of gradient components in geological background

The case where the parameter estimation using all six gradient components is performed simultaneously in the full geological background plus $1E^2/\text{Hz}$ white noise is tested. The a priori information of (7.3) and (7.4) with $\sigma_D^2 = 100\text{m}$ is used. The physical correlations between observations are taken into account; the equation of (A.1) and the instrumental noise N_F in Table A.1 is used.

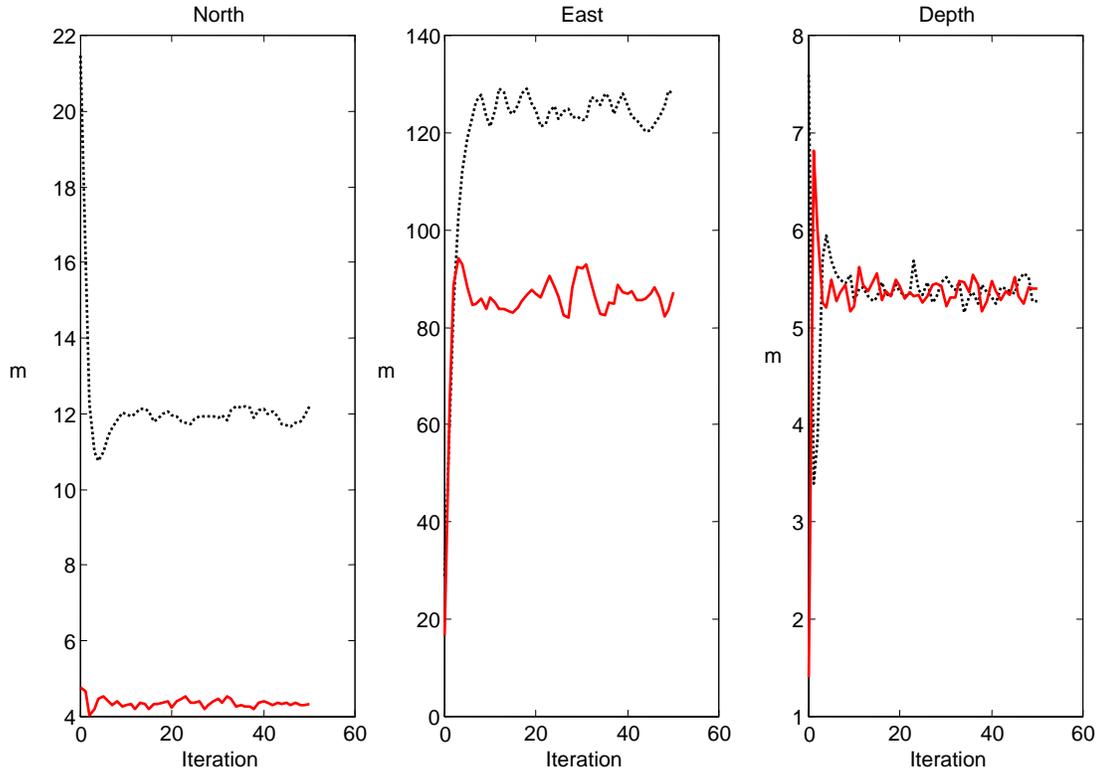


Figure 7.15: CASE VII: The absolute errors of location; black-dotted line indicates a geological model was not removed; red-solid line indicates a geological model was removed.

Figure 7.15 shows the results of all combined gradients. When six gradient components are simultaneously observed, the absolute errors of north and east coordinates for the case of WOB (Psd model 2 was used) (red-solid line) are smaller than those for the case of WB (Psd model 1 was used) (black-dotted line). For instance, the north coordinate error of the latter case is about 12m while it is 3m for the former case at

iteration 50. The mean square prediction errors, $\text{MSPE}\{\tilde{\Theta}_{50}\}$, obtained from (7.2) for both cases at iteration 50 are:

(a) The geological model was not removed (Psd model 1 was used) and $\hat{\sigma}_0^2 = 68.181258$:

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 68.181258 \begin{bmatrix} 11.921697 & 0.000437 & 0 \\ 0.000437 & 339.275122 & 0 \\ 0 & 0 & 10.570163 \end{bmatrix} [\text{m}^2] \quad (7.17)$$

(b) The geological model was removed (Psd model 2 was used) and $\hat{\sigma}_0^2 = 33.682292$:

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = 33.682292 \begin{bmatrix} 12.637481 & 0.000855 & 0 \\ 0.000855 & 341.884744 & 0 \\ 0 & 0 & 11.131837 \end{bmatrix} [\text{m}^2] \quad (7.18)$$

Note that the diagonal elements of (7.17) are slightly smaller than those of (7.18). This is due to the influence of psd modeling as discussed in CASE VI. Finally, the root-mean-square-error's of the adjusted parameters taken from the diagonal elements of (7.17) and (7.18) with $\hat{\sigma}_0^2$ are given in Table 7.15a. Table 7.15b shows the hypothesis testing at the significance level $\alpha = 0.05$ after 50 iterations. Using F-test, the null hypotheses for both cases are accepted.

For the χ^2 test, the rejection of the null hypothesis (two-tailed test) indicates an incomplete reduction of the noise background in the observation \mathbf{y} in eq. (6.13).

	rmse(m)		
	Northing	Easting	Depth
WB	28.510	152.092	26.845
WOB	20.631	107.310	19.363

(a)

	Test statistic	
	WB	WOB
A	0.007862339	0.004185320
B	68.181471318	33.682397807
T	0.000115314	0.000124258
$T_\chi \times 10^{-5}$	654.540	323.350

$T = \frac{A}{B}$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$ eq. (6.27)
 $\psi_0 = F_{1-\alpha}(m, n - m) = 2.604918546$ eq. (6.28)
 Reject H_0 if $T_\chi > \chi^2_{1-\frac{\alpha}{2}, n-m} = 9.627E+5$ or $T_\chi < \chi^2_{\frac{\alpha}{2}, n-m} = 9.572E+5$
 with $T_\chi = (n - m)\bar{\sigma}_0^2 / \sigma_0^2 = \mathbf{\Omega} / \sigma_0^2$, $\alpha = 0.05$, $m = 3$, $n = 6 \times 400 \times 400$, and $\sigma_0^2 = 1$
 WB = full geological background plus $1E^2/\text{Hz}$ white noise
 WOB = full geological background plus $1E^2/\text{Hz}$ white noise minus a geological model

(b)

Table 7.15: CASE VII: (a) The root-mean-square-error (rmse) of the adjusted parameters after 50 iterations; (b) Hypothesis testing via F-test and χ^2 test after 50 iterations.

7.9 CASE VIII: Predicted location and orientation in $0.01E^2/\text{Hz}$ white noise environment using six gradient components

For this case, we want to evaluate how well the parameter estimation using the random effects model approach is able to refine the azimuthal orientation of the $2\text{m} \times 2\text{m} \times 1000\text{m}$ anomaly. Suppose the anomaly is rotated by 47 degrees with respect to north direction in the $0.01E^2/\text{Hz}$ white noise environment. Let the output of the matched filtering be given as in the previous cases except that the azimuthal orientation will be determined using six gradient components. We shall also determine how the hypothesis testing in association with the choices of σ_D^2 behaves when estimation of the orientation is included. Here, we consider no correlation between sensors under white noise circumstances. Provided that the approximate orientation is obtained from the detection process, we rewrite (7.3) and (7.4):

$$\boldsymbol{\theta}_0 = \begin{bmatrix} \text{North} \\ \text{East} \\ \text{Depth} \\ \beta' \end{bmatrix} = \begin{bmatrix} 21030.00\text{m} \\ 18000.00\text{m} \\ 60.00 \\ 50^\circ \end{bmatrix} \quad (7.19)$$

together with the a priori covariance matrix

$$\boldsymbol{\Sigma}_\theta = \sigma_0^2 \begin{bmatrix} \sigma_N^2 & 0 & 0 & 0 \\ 0 & \sigma_E^2 & 0 & 0 \\ 0 & 0 & \sigma_D^2 & 0 \\ 0 & 0 & 0 & \sigma_{\beta'}^2 \end{bmatrix} \quad \begin{array}{l} \sigma_N = \pm 30.00\text{m} \\ \text{with } \sigma_E = \pm 30.00\text{m} \\ \sigma_{\beta'} = \pm 3^\circ \end{array} \quad (7.20)$$

(a) $\sigma_D = \pm 1.00\text{m}$, and (b) $\sigma_D = \pm 10.00\text{m}$

Figure 7.16a shows that the absolute errors of parameter estimates using $\sigma_D = \pm 1.00\text{m}$ and $\sigma_D = \pm 10.00\text{m}$ are slightly different. The estimation method can refine the orientation to better than 0.172° after 50 iterations. The mean square prediction errors, $\text{MSPE}\{\tilde{\boldsymbol{\Theta}}_{50}\}$, after 50 iterations obtained from (7.2) are as follows:

(i) For the case of $\sigma_D^2 = 1\text{m}^2$ and $\hat{\sigma}_0^2 = 0.998844$,

$$\text{MSPE}\{\tilde{\boldsymbol{\Theta}}_{50}\} = \hat{\sigma}_0^2 \begin{bmatrix} 4.073396\text{m}^2 & 4.124023\text{m}^2 & -2.664\text{E} - 7\text{m}^2 & -0.000034\text{m} \cdot \text{deg} \\ 4.124023\text{m}^2 & 4.643856\text{m}^2 & -3.267\text{E} - 7\text{m}^2 & -0.000041\text{m} \cdot \text{deg} \\ -2.664\text{E} - 7\text{m}^2 & -3.267\text{E} - 7\text{m}^2 & 0.179762\text{m}^2 & 0.008235\text{m} \cdot \text{deg} \\ -0.000034\text{m} \cdot \text{deg} & -0.000041\text{m} \cdot \text{deg} & 0.008235\text{m} \cdot \text{deg} & 1.051294\text{deg}^2 \end{bmatrix} \quad (7.21)$$

(ii) For the case of $\sigma_D^2 = 100\text{m}^2$ and $\hat{\sigma}_0^2 = 0.998845$,

$$\text{MSPE}\{\tilde{\Theta}_{50}\} = \hat{\sigma}_0^2 \begin{bmatrix} 4.072151\text{m}^2 & 4.122719\text{m}^2 & 4.516\text{E} - 7\text{m}^2 & -0.000046\text{m} \cdot \text{deg} \\ 4.122719\text{m}^2 & 4.642434\text{m}^2 & -3.390\text{E} - 7\text{m}^2 & -0.000035\text{m} \cdot \text{deg} \\ 4.516\text{E} - 07\text{m}^2 & -3.390\text{E} - 7\text{m}^2 & 0.218655\text{m}^2 & 0.010155\text{m} \cdot \text{deg} \\ -0.000046\text{m} \cdot \text{deg} & -0.000035\text{m} \cdot \text{deg} & 0.010155\text{m} \cdot \text{deg} & 1.051021\text{deg}^2 \end{bmatrix} \quad (7.22)$$

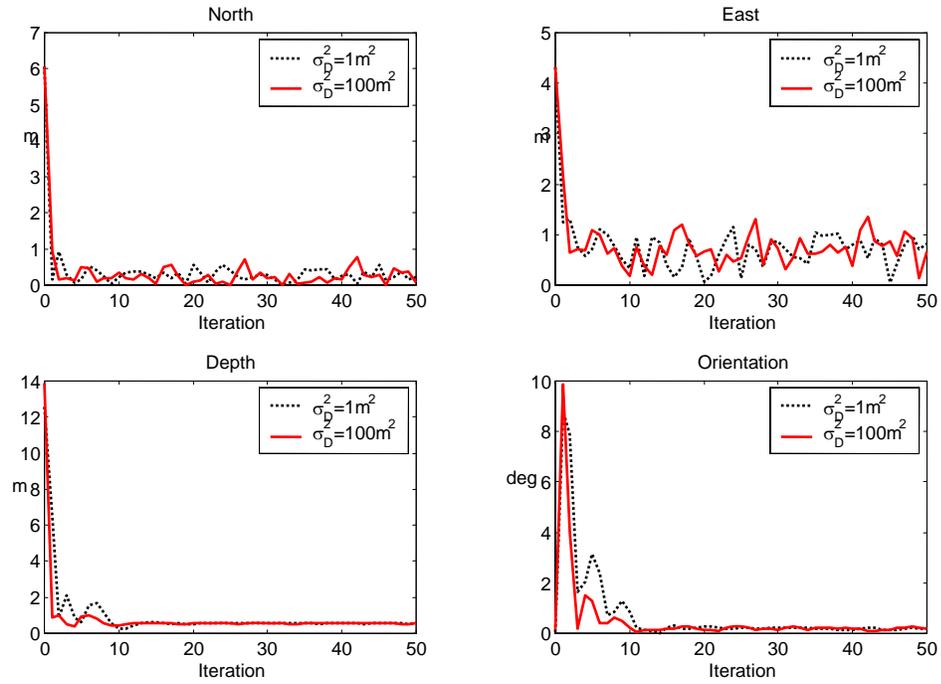
As seen above, the depth and orientation estimates are slightly correlated. Finally, the rmse's of the adjusted parameters taken from (7.21) and (7.22) are summarized in Table 7.16. In Figure 7.16b, the test statistic for the case of $\sigma_D^2 = 100\text{m}^2$ decreases faster than the other; and, in turn, is below the computed threshold ψ_0 for the specified significance level, $\alpha = 0.05$ after 3 iterations. The hypothesis tests for both cases at iteration 50 are summarized in Table 7.17.

Choice	rmse			
	Northing (m)	Easting (m)	Depth (m)	Orientation (deg)
$\sigma_D^2 = 1\text{m}^2$	2.016792	2.153386	0.467336	1.024601
$\sigma_D^2 = 100\text{m}^2$	2.017099	2.153715	0.423738	1.024733

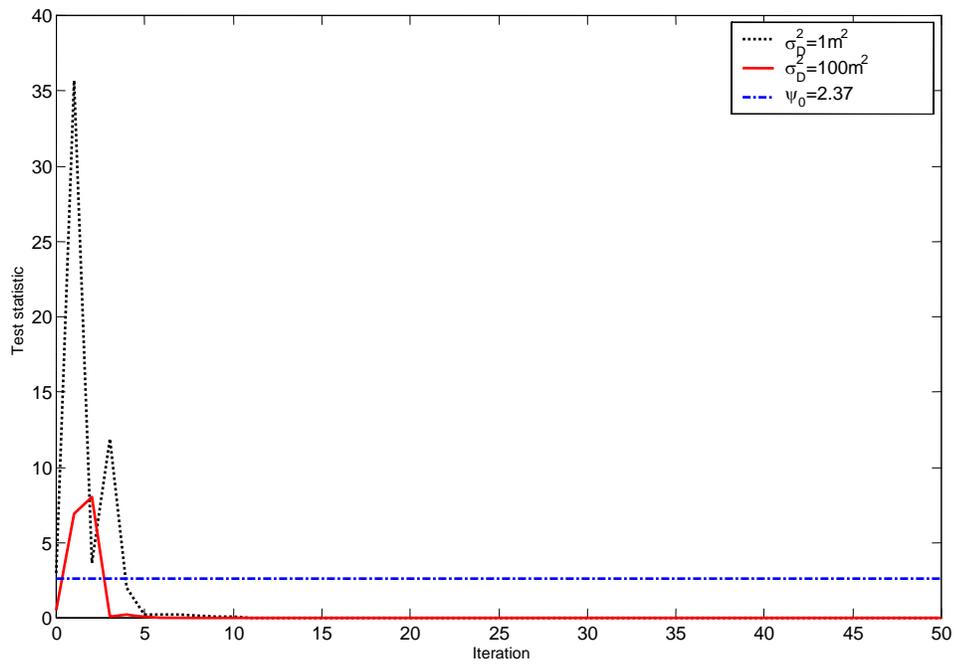
Table 7.16: CASE VIII: The root-mean-square-errors (rmse's) of the adjusted parameters after 50 iterations.

Choice	Test statistic		
	A	B	T
$\sigma_D^2 = 1\text{m}^2$	0.000271891	0.998848584	0.000272205
$\sigma_D^2 = 100\text{m}^2$	0.000913490	0.998849225	0.000914543
$T = \frac{A}{B}$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$			eq. (6.27)
$\psi_0 = F_{1-\alpha}(m, n - m) = 2.371941509$ with $\alpha = 0.05$, $m = 4$, and $n = 6 \times 400 \times 400$			eq. (6.28)

Table 7.17: CASEVIII: Hypothesis testing after 50 iterations.



(a)



(b)

Figure 7.16: CASEVIII: (a) The absolute errors of location and orientation; (b) Test statistic

We further evaluate the robustness of the random effects model approach in estimating orientation by adding either $1E^2/\text{Hz}$ white noise alone or the full geological background plus $1E^2/\text{Hz}$ white noise to the anomaly's signal. Unfortunately, neither case indicates that the parameter estimation can be done successfully. The adjusted solutions of the orientation fluctuate and are unstable in the iterative procedure. Furthermore, the location errors (north and east coordinates and depth) continuously grow with respect to the number of iterations. If we estimate the orientation alone by assuming the location is known perfectly [by letting $\sigma_N^2 \rightarrow 0$, $\sigma_E^2 \rightarrow 0$, and $\sigma_D^2 \rightarrow 0$, (i.e., very large weights are applied)], the adjusted solutions still diverge (not shown here). The test statistic does not give any reliable judgment; H_0 is accepted although there is a large discrepancy between the adjusted estimates and the a priori values Θ_0 . As such, to estimate the location as well as the orientation parameters, it requires a complete reduction of the observations; any noisy contents should be removed as much as possible.

7.10 CASE IX: Estimation of density contrast

For the simplicity of density contrast estimation, suppose the location and the orientation of the 2m x 2m x 1000m anomaly are perfectly known; see the expressions (B.1) and (B.2) in Appendix B. We wish to estimate the density contrast $\Delta\rho$ alone.

Let the true value of the density contrast be -1.500 g/cm^3 . We want to evaluate the robustness of the random effects model approach when estimating $\Delta\rho$. Given the a priori information, $\theta_0 = -1.400 \text{ g/cm}^3$ with $\sigma_{\Delta\rho}^2 = 0.01 \text{ [g/cm}^3\text{]}^2$, we determine the density contrast in different noise backgrounds. All six combined gradients are employed in this case. Because of the linear estimation according to (B.1) and (B.2) (i.e., assuming the other parameters are completely known), no iterative procedure is required. Consider no correlations between sensors if white noise is used.

The estimation results are summarized in Table 7.18. The estimate of the density contrast, $\Delta\hat{\rho}$, agrees within less than 0.020 g/cm^3 (see its rmse obtained from (7.2)) with the true value of density contrast for the case of $0.01 \text{ E}^2/\text{Hz}$ white noise. However, the estimation errors become larger as stronger noise content is added to the anomaly's gradient signals. For the extreme case, where the full geological background plus $1 \text{ E}^2/\text{Hz}$ white noise is used, the estimate deviates from the true value by 1.009 g/cm^3 . As such, the F-test fails to accept H_0 at the 0.05 level of significance. For the density estimation, the rejection of H_0 does not mean that the estimate is incorrect. Rather, it implies that the estimate and the a priori Θ_0 do not coincide in the statistical sense of (6.4). Is the choice of Θ_0 suitable for the adjustment system? or Does the noise background greatly affect the estimation? As such, the χ^2 test on errors in the stochastic model and/or due to an incomplete reduction of the observations is applied. As listed in Table 7.18, the χ^2 test fails to accept H_0 on the upper tail for the case of full geological background plus $1 \text{ E}^2/\text{Hz}$ white noise. The test statistic is large (654.531×10^5), relative to the two-tail criterion. However, when the noise background decreases, for $0.01 \text{ E}^2/\text{Hz}$ white noise case, the test statistic T_χ falls inside the acceptance interval.

Noise	$\Delta\hat{\rho}$ (g/cm ³)	rmse (g/cm ³)	F-test			χ^2 -test
			A	B	T	$T_\chi \times 10^{-5}$
0.01E ² /Hz	-1.507	0.023	1.219119350	0.998843524	1.220530864	9.589
1E ² /Hz	-1.387	0.092	0.105302483	1.000662128	0.105302483	9.606
WOB	-1.832	0.535	126.059658947	33.682196525	3.742619898	323.349
WB	-2.509	0.763	839.760390702	68.180358563	12.316749404	654.531
$T = \frac{A}{B}$ with $A = \mathbf{R}/m$ and $B = \mathbf{\Omega}/n - m$						eq. (6.27)
$\psi_0 = F_{1-\alpha}(m, n - m) = 3.841468506$						eq. (6.28)
Reject H_0 if $T_\chi > \chi^2_{1-\frac{\alpha}{2}, n-m} = 9.627E+5$ or $T_\chi < \chi^2_{\frac{\alpha}{2}, n-m} = 9.573E+5$						
with $T_\chi = (n - m)\bar{\sigma}_0^2 / \sigma_0^2 = \mathbf{\Omega} / \sigma_0^2$, $\alpha = 0.05$, $m = 1$, $n = 6 \times 400 \times 400$, and $\sigma_0^2 = 1$						
WB = full geological background plus 1E ² /Hz white noise						
WOB = full geological background plus 1E ² /Hz white noise minus a geological model						

Table 7.18: CASE IX: The estimates of density contrast and hypothesis testing

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

In this study, the two dimensional matched filters for six gradient components observed at low aircraft altitude was developed and tested in white and correlated noise environments for shallow small mass anomalous detection. The airborne gradiometric data was simulated using the 1987 GGSS data, 30m DEM, and $1E/\sqrt{\text{Hz}}$ white Gaussian noise. Different sizes of a rectangular box representing the true anomalies were simulated and these signals were added to the simulated noise. The reciprocal psd models were applied to the matched filters for different types of geological backgrounds: (1) full geological background plus $1E^2/\text{Hz}$ white noise, (2) geological background plus $1E^2/\text{Hz}$ white noise minus a geological model, and (3) the GGSS data plus $1E^2/\text{Hz}$ white noise. For comparison, the matched filters using the empirical psd's obtained from the geological backgrounds were also considered.

Numerical analyses of the capability of the matched filters were presented. The results in Chapter 5 indicate that, overall, the Γ_{33} gradient component is able to detect the anomalies of interest using matched filters in this simulation. Psd modeling affects the detectability of matched filters, where the empirical psd's generally yield superior results. Although more accurate psd models are needed to enhance the capability of matched filters, the matched filter for all six combined gradients shows an improvement of detection superior to the use of matched filters for individual gradients. The correlated noise environments slightly affect the detectability of the filters using the empirical psd's. For the best performance, the knowledge of the signal waveform is required as well as accurate psd's of the background noise.

By applying the matched filter to the simulation observations including the anomaly signal to be detected, the highest output peak indicates the likely location of the signal. Statistically, if the peak exceeds the threshold for a given POM (or a significance level), the signal is presumed to be present at the location of the peak. The threshold with a computed POF does not in any way influence the selection of the candidate location; though, it sets up a level above which all values of the output, maximum or not, have a probability of being a true location. In this sense, the threshold does not give a good measure of the probability of success since it does not account for the maximum value of the output being special, it rather allows possible candidate(s) of the target anomaly. However, since a matched filter maximizes the SNR, the POF is useful when making a decision concerning the presence of the target signal. The choice of POM should not be too high, otherwise until the corresponding value of POF may not be reliable; increasing POM decreases POF. In this study, we find that $\text{POM} = 0.001$ is appropriate.

In the case of full geological background plus $1E^2/\text{Hz}$ white noise minus the geological model computed using 120m resolution topography, the matched filters performed not better than in the case of full geological background plus $1E^2/\text{Hz}$ white noise. This lack of improvement despite reduction in background noise was traced to high sensitivity of the filter to the psd models used. In this case, Psd model 2 may not adequately characterize the noise background.

In comparison with the use of psd models, we used the empirical psd's and applied them for different cases of geological backgrounds. We found that geological background slightly affects the detectability performance of the matched filters. Another factor is the magnitudes of noise, which may be large, compared to the gradient signals of the target anomaly. As a consequence, the target signal may not be detected by the filters. Numerical tests in Table 5.20 show that the locations of target anomalies within the background field also affect the detectability of the matched filters.

The target anomaly located on observation grid intersections can be detected better than the anomaly located within a grid square. We found that with higher resolution in the observations the detectability of the matched filters increases. The detectability of the filters is sensitive to orientations. When the filter is rotated by a few degrees with respect to the true orientation of the target anomaly, it can not detect the anomaly successfully. A bank of filters at different azimuthal orientations is required to detect a signal with unknown orientation. The performance of the matched filters for different cases was summarized in the plots in Section 5.7.

If the matched filter output provides a good approximation of the true location of the anomaly, parameter estimation theory (developed in Chapter 6) may be used to refine the location, including its depth. To refine the detection part, as well as determine additional parameters that were not included in the detection, we use the random effects model in a least-squares adjustment procedure, which allows a priori information of a stochastic nature on the parameters. Using individual gradient components, overall, the component Γ_{33} provides an improvement of location estimates in different background noises. Moreover, estimation can be improved using the six-gradient combination, not only in the cases of $0.01E^2/\text{Hz}$ and $1E^2/\text{Hz}$ white noises, but also in the case where a geological model was subtracted from the full geological background plus $1E^2/\text{Hz}$ white noise. However, the result of estimating the location is worst for the case of the full geological background plus $1E^2/\text{Hz}$ white noise, both when using individual gradients or all six gradients.

In our study, since we deal with multiple data tracks and single or multiple sensors, the inversion of the corresponding large covariance matrix may be numerically difficult. We have applied an orthogonal transformation to the random effects model so that the transformed covariance matrix becomes diagonal (or block diagonal), and is easily inverted (assuming we have regularly gridded data).

The test statistic (or significance test) which follows the F-distribution has been developed in order to determine whether the adjusted parameters give us misleading results on the basis of the a priori information given by the matched filter. However, the test does not appear to be useful in assessing the estimation performance in the presence of strong noise backgrounds.

Nonetheless, the random effects models have shown to be successful in refining the location and orientation using all six combined gradients--the adjusted values are close to the true parameters with significantly reduced errors (from the matched filter result). However, the adjusted solutions (iterated due to the nonlinearity of the model) do not converge if only individual gradients are used in the adjustment (with full geological background plus $1\text{E}^2/\text{Hz}$ white noise). The introduction and implementation of an additional χ^2 test confirms that an incomplete reduction of observation noise and an inaccurate psd model were applied.

An additional parameter, the density contrast, was considered. Again, both F-test and χ^2 test failed for the computed thresholds based on $\alpha = 0.050$. More accurate psd models are needed as well as a requirement of removing correlated noise contents in the observations as much as possible.

In this study, we have considered only one anomaly to be detected. Multiple targets should be the focus of future work. These problems can make the detection and estimation approach much more difficult. Multiple hypothesis testing would be recommended to apply to the approach in order to classify the targets of interest. Furthermore, the anomalies may have densities, slightly different from the geological background. This problem can make detecting the signal and the estimation of unknown parameters much more difficult. A more accurate and informative geological model or additional sensors (e.g., ground penetrating radar or electro magnetics) may be needed.

Finally, much more work with real data would be recommended to test the robustness of the signal detection and parameter estimation approach proposed in this study to further refine its practical applications in concert with the use of a future gradiometer with $1\text{E}^2/\text{Hz}$ instrumental noise sensitivity.

REFERENCES

- Brigham, E.O. (1988): *The Fast Fourier Transform (FFT) and its Applications*. Prentice Hall, NJ.
- Cadzow, J.A. (1987): *Foundations of Digital Signal Processing and Data Analysis*. MacMillan Pub. Co., New York.
- Chen, C. and S.A. Kassam (1985): Robust multiple-input matched filtering: frequency and time-domain results. *IEEE Trans. on Info. Theory*, IT-31(6), 812-821.
- Chinnery, M.A. (1961): Terrain corrections for airborne gravity gradient measurements. *Geophysics*, 26(4), 480-489.
- DiFranco, J.V. (1968): *Radar Detection*, Prentice-Hall, Inc., NJ.
- Dorman, L.M. and B.T.R. Lewes (1974): The use of nonlinear functional expansions in calculation of the terrain effect in airborne and marine gravimetry and gradiometry. *Geophysics*, 39(1), 33-38.
- Dransfield, M.H. (1994): Airborne gravity gradiometry. Ph.D. Thesis, Department of Physics, University of Western Australia.
- Dumrongchai, P. and C. Jekeli (2004a): On detecting shallow mass anomalies with ground and airborne gradiometry by filtering techniques. Presented at 2004 Joint Assembly: CGU, AGU, SEG, and EEGS, Montreal, Canada, May 17-21, 2004.
- Dumrongchai, P. and C. Jekeli (2004b): Small mass anomaly detection from airborne gradiometry by a combination of matched filter and hypothesis testing. Presented at the International Symposium on Gravity, Geoid, and Space Missions, Porto, Portugal, Aug. 30 - Sept. 3, 2004.
- Forsberg, R. (1984): A study of terrain reductions, density anomalies and geophysical inversion methods in gravity field modeling. Department of Geodetic Science and Surveying, Report No. 355, The Ohio State University, Columbus, Ohio.
- Geraniotis, E. and H.V. Poor (1987): Robust matched filters for optical receivers. SRC TR 87-154, Department of Electrical Engineering, University of Maryland, College Park, Maryland.

- Gleason, D.M. (1988): Computing any arbitrary downward continuation kernel function of an integral predictor yielding surface gravity disturbance components from airborne gradient data. *Manuscripta geodaetica*, 13, 147-155.
- Gunn, P.J. (1975): Detection of bodies with specified density-magnetization ratios by the use of multichannel Wiener filters. *Geophysics*, 40(4), 617-620.
- Gunn, P.J. (1978): Inversion of airborne radiometric data. *Geophysics*, 43(1), 133-143.
- Gupta, V.K. and N. Ramani (1980): Some aspects of regional-residual separation of gravity anomalies in a Precambrian terrain. *Geophysics*, 45(9), 1412-1426.
- Hammer, S. (1976): Topographic corrections for gradients in airborne gravimetry. *Geophysics*, 41(6), 1346-1352.
- Heiskanen, W.A. and H. Moritz (1979): *Physical Geodesy*. San Francisco (Reprint 1996 by Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, Columbus, Ohio).
- Helstrom, C.W. (1960): *Statistical theory of signal detection*. Pergamon Press.
- Helstrom, C.W. (1995): *Elements of signal detection and estimation*, PTR Prentice Hall.
- Hofmann-Wellenhof and H. Moritz. (2005): *Physical geodesy*. 2nd edition. Springer.
- Jekeli, C. (1985): On optimal estimation of gravity from gravity gradients at aircraft altitude. *Rev. of Geophys.*, 23(3), 301-311.
- Jekeli, C. (1986): Estimation of gravity differences from a large and densely spaced heterogeneous gradient data set using an integral formula. *Manuscripta geodaetica*, 11, 48-56.
- Jekeli, C. (1987a): New instrumentation techniques in geodesy. *Rev. Geophysics*, 25, 889.
- Jekeli, C. (1987b): The downward continuation of aerial gravimetric data without density Hypothesis. *Bulletin Géodésique*, 61, 319-329.
- Jekeli, C. (1993): A review of gravity gradiometer survey system data analyses. *Geophysics*, 58(4), 508-514.
- Jekeli, C. (2000): *Inertial navigation systems with geodetic applications*. Walter deGruyter, Berlin.

- Jekeli, C. (2003): Statistical analysis of moving-base gravimetry and gravity gradiometry. Reports of the Department of Civil and Env. Eng. and Geodetic Science, Report No. 466, The Ohio State University.
- Jekeli, C. (2006): Decision criteria for signal detection and estimation. Presented at Gradiometry workshop, National Geospatial-Intelligence Agency (NGA), March 14-15, St. Louis, MO.
- Jekeli, C. and L. Zhu (2006): Comparison of methods to model gravitational gradients from topographic data bases. *Geophysical Journal International*, 166(3), 999 - 1014.
- Jekeli, C., P. Dumrongchai, and L. Zhu (2003): Analytic formulas for the gravitational gradients generated by triangular vertical prisms with sloped tops. Presented at IUGG2003, IAG Symposium, June 11 – July 11, Sopporo, Japan.
- Jordan, S.K. (1978): Statistical model for gravity, topography, and density contrasts in the Earth. *Journal of Geophysical Research*, 83(B4), 1816-1824.
- Kasevich, M. and S. Chu (1991): Atomic interferometry using stimulated Raman transitions. *Phys. Rev. Lett.*, 67, 181-184.
- Kastella, K. (2003): Gravity gradiometry signal processing for UGF detection and Imaging, Presented by Veridian Systems and Jet Propulsion Laboratory (JPL) at JPL, California Institute of Technology in Feb., 2003.
- Kay, S.M. (1993): *Fundamentals of statistical signal processing Vol. I: estimation theory*. PTR Prentice Hall, NJ.
- Kay, S.M. (1998): *Fundamentals of statistical signal processing Vol. II: detection theory*. PTR Prentice Hall, NJ.
- Klinge, E.E., I. Marson, and H.-G. Kahle (1991): Automatic interpretation of gravity gradiometric in two dimensions: vertical gradient. *Geophysical Prospecting*, 39, 407-434.
- Koch, K.R. (1999): *Parameter estimation and hypothesis testing in linear models, 2nd edition*. Springer-Verlag, NY.
- Lancaster-Jones, E. (1932): Principles and practice of the gravity gradiometer, *J. Sci. Instrum.*, 9, 491-506.
- Lee, J.B. (2001): FALCON gravity gradiometer technology. Proceedings of the Gradiometry Workshop. The 71 SEG International Exposition, 2001.

- Leick, A (1995): *GPS satellite surveying, 2nd edition*. John Wiley & Sons, INC., NY.
- Lemoine, F.G., S.C. Kenyon, J.K. Factor, R.G. Trimmer, N.K. Pavlis, D.S. Chinn, C.M. Cox, S.M. Klosko, S.B. Luthcke, M.H. Torrence, Y.M. Wang, R.G. Williamson, E.C. Pavlis, R.H. Rapp, and T.R. Olson (1998): The development of the joint NASA GSFC and the National Imagery and Mapping Agency (NIMA) Geopotential Model EGM96. NASA/TP-1998-206861.
- Lorenzini, E.C., G.E. Gullahorn, F. Fuligni (1988): Recent developments in gradiometry from the space-shuttle-borne tethered satellite system. *J. Appl. Phys.*, 63, 216-223.
- Lui, G., P. Diorio, P. Stone, G. Lockhart, A. Christensen, N. Fitton, and M. Dransfield (2001): Detecting kimberlite pipes at Ekati with airborne gravity gradiometry. Extended Abstracts, ASEG 15th, Geophysical Conference and Exhibition, August 2001, Brisbane.
- Marson, I. and E.E. Klingele (1993): Advantages of using the vertical gradient of gravity for 3-D interpretation. *Geophysics*, 58(11), 1588-1595.
- Matthews, R. (2002): Mobile gravity gradiometry. Ph.D. Thesis, Department of Physics, University of Western Australia.
- McCarthy, D.D. (1996): IERS Technical Note 21. www.iers.org.
- McGuirk, J.M., G.T. Foster, J.B. Fixler, M.J. Snadden, and M.A. Kasevich (2002): Sensitive absolute-gravity gradiometry using atom interferometry. *Phys. Rev. A*, 65, 033608.
- Middleton, D. (1960a): *An introduction to statistic communication theory*. McGraw-Hill Book Company, NY.
- Middleton, D. (1960b): On new classes of matched filters and generalizations of the matched filter concept. *IRE Trans. Information Theory*, IT-6, 349-360.
- Middleton, D. (1965): *Topics in communication theory*. McGraw-Hill Book Company, NY.
- Mikhail, E.M. and F. Ackerman (1976): *Observations and least squares*. Harper & Row, Publishers, NY.
- Moritz, H. (1980): *Advanced Physical Geodesy*. Herbert Wichman Verlag, Karlsruhe (Reprint 2001 by Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, Columbus, Ohio).

- Nettleton, L.L. (1976): *Gravity and Magnetism in Oil Prospecting*. MacGraw Hill Book Co., New York.
- North, D.O. (1943): Analysis of the factors which determine signal/noise discrimination in radar, RCA Laboratories, Princeton, N.J., Rept. PTR-6C.
- Paik, H.J. and M.V. Moody (1993): Gauss's law test of gravity at short range. *Physical Review Letters*, 70, 1195-1198.
- Parker, R.L. (1972): The rapid calculation of potential anomalies. *Geophys. J.R. Astron. Soc.*, 31, 447-455.
- Parker, R.L. (1994): *Geophysical Inverse Theory*. Princeton University Press, N.J.
- Pawlowski, R.S. and R.O. Hansen (1990): Gravity anomaly separation by Wiener filtering. *Geophysics*, 55(5), 539-548.
- Peters, A., K.Y. Chung, and S. Chu (2001): High-precision gravity measurements using atom interferometry. *Metrologia*, 38, 25-61.
- Poor, H.V. (1983): Robust matched filters. *IEEE Trans. on Info. Theory*, IT-29(5), 677 - 687.
- Romaides, A.J., J.C. Battis, R.W. Sands, A. Zorn, D.O. Benson Jr., and D.J. DiFrancesco (2001): A comparison of gravimetric techniques for measuring subsurface void signals. *J. Phys. D: Appl. Phys.*, 34, 433-443.
- Schaffrin, B. (1985): A note on linear prediction within a Gauss-Markoff model linearized with respect to a random approximation. Proc. First Tampere Sem. Dept. of Math. Sci., Univ. Tampere, Finland, 285-300.
- Schaffrin, B. (1987): Less sensitive tests by introducing stochastic linear hypotheses. Proc. Second Int. Tampere Conf. in Stat., Tampere, Finland, 1-4 June 1987, 647-664.
- Schaffrin, B. (1989): Tests on random effects based on homogeneously linear predictors. Proc. Int. workshop on "Theory and Practice in Data Analysis", 19-21 August 1988, Berlin (East), Report No. R-MATH-01/89, 209-227.
- Schaffrin, B. (2001): On testing softly unbiased kriging results on the sphere. Proc. Geostats Congr. 2000 Cape Town. Geostatist Assoc. of Southern Africa, Pretoria, 195-204.
- Schaffrin, B. and Y. Bock (1994): Geodetic deformation analysis based on robust inverse theory. *Manus. Geodaet.*, 19, 31-44.

- Smith, D.A., D.S. Robertson, and D.G. Milbert (2001): Gravitational attraction of local crustal masses in spherical coordinates. *Journal of Geodesy*, 74, 783-795.
- Smith, R.S., J.B. Thurston, T.F. Dan, and I.N. MacLeod (1998): iSPITM-the improved source parameter imaging method. *Geophysical Prospecting*, 46, 141-151.
- Snadden, M.J., J.M. McGuirk, P. Bouyer, K.G. Haritos, and M.A. Hasevich (1998): Measurement of the earth's gravity gradient with an atom interferometer-based gravity gradiometer. *Phys. Rev. Lett.*, 81(5), 971-974.
- Streland, A.H. (2003): Going deep: a system concept for detecting deeply buried facilities from space. Thesis, Air War college, Air University, U.S.A.
- Talwani, M. (2003): Joint inversion and other issues concerning gravity gradients. <http://cohesion.rice.edu/naturalsciences/earthscience/index.cfm>.
- Thomson, D.T. (1982): EULDPH: A new technique for making computer-assisted depth estimates from magnetic data. *Geophysics*, 47(1), 31-37.
- Thurston, J.B. and R.S. Smith (1997): Automatic conversion of magnetic data to depth, dip, and susceptibility contrast using the SPI(TM) method. *Geophysics*, 62(3), 807-813.
- Torge, W. (1989): *Gravimetry*. Walter deGruyter, Berlin.
- Torge, W. (2001): *Geodesy*. 3rd edition. Walter deGruyter, Berlin.
- Turin, G.L. (1960): An introduction to matched filters. *IRE Trans. Information Theory*, IT-6, 310-329.
- Turin, G.L. (1976): An introduction to digital matched filters. *Proceedings of the IEEE*, 64(7), 1092-1112.
- Tziavos, I.N., M.G. Sideris, R. Forsberg, and K.P. Schwarz (1988): The effect of the terrain on airborne gravity and gradiometry. *J. of Geophys. Res.*, 93(B8), 9173-9186.
- van Kann, F.J. (1992): End-of-Grant Report: Project No. 880, ERDC, Department of Industry and Tecnology, Canberra.
- van Kann, F.J., M.J. Buckingham, M.H. Dransfield, C. Edwards, A.G. Mann, R. Matthews, and P.J. Turner (1990): Laboratory tests of a superconducting gravity gradiometer. Proceedings of the 19th International Conference on Low Temperature Physics, *Physica B*, 165 & 166, 93-94.

- van Leeuwen, E.H. (2000): BHP develops airborne gravity gradiometer for mineral exploration, *The Leading Edge*, 19(12), 1296-1297.
- Van Trees, H.L. (1968): *Detection, Estimation, and Modulation Theory: Part I*. John Wiley and Son, Inc., NY.
- Vasco, D.W. (1989): Resolution and variance operators of gravity and gravity gradiometry. *Geophysics*, 54(7), 889-899.
- Vasco, D.W. and C. Taylor (1991): Inversion of airborne gravity gradient data, southwestern Oklahoma. *Geophysics*, 56(1), 90-101.
- Wang, Y.M. (1988): Determination of the gravity disturbance on the Earth's topographic surface from airborne gravity gradient data. Department of Geodetic Science and Surveying, Report No. 401, The Ohio State University, Columbus, Ohio.
- Whalen, A.D. (1971): *Detection of Signals in Noise*. Academic Press, NY.
- White, J.V. and J.D. Goldstein (1984): Gravity gradiometer survey data processing, AGFL-TR-84-0198, report prepared by The Analytic Sciences Corporation for the Air Force Geophysics Laboratory, Hanscom AFB.
- White, J.V., Sailor, R.V., Lazarewicz, A.R., and Leschack, A.R. (1983): Detection of seamount signatures in SEASAT altimeter data using matched filters. *J. Geophys. Res.*, 88, 1541-1551.
- Wiener, N (1949): *Extrapolation, interpolation and smoothing of stationary time series*. Wiley and Sons, N.Y.
- Zhdanov, M.S., R. Ellis, and S. Mukherjee (2004): Three-dimensional regularized focusing inversion of gravity gradient tensor component data. *Geophysics*, 69(4), 925-937.

APPENDIX A

THE RECIPROCAL DISTANCE POWER SPECTRAL DENSITY MODELS OF GRVITATIONAL GRADIENTS

The reciprocal (auto or cross) psd models of gravitational gradients are given by (Jekeli, 2003):

$$\mathbf{P}_{j_1 j_2, k_1 k_2} = (\lambda_{j_1})(\lambda_{j_2})(\lambda_{k_1})^*(\lambda_{k_2})^* \Phi_T \text{ with } j_1, j_2, k_1, k_2 \in \{1, 2, 3\} \quad (\text{A.1})$$

where $\lambda_m = i2\pi f_m$ for $m = 1, 2$ and $\lambda_3 = 2\pi f$ with $f = \sqrt{f_1^2 + f_2^2}$. The psd of disturbing potential Φ_T is given by:

$$\Phi_T(f_1, f_2, x_3, x'_3) = \left(\sum_{j=0}^J \frac{\sigma_j^2}{\alpha_j f} e^{-\frac{2\pi f}{\alpha_j}} \right) e^{-2\pi f(x_3 + x'_3)} \quad (\text{A.2})$$

J=16	PSD MODEL 1		PSD MODEL 2		PSD MODEL 3	
j	σ^2 (m ⁴ /s ⁴)	α (1/m)	σ^2 (m ⁴ /s ⁴)	α (1/m)	σ^2 (m ⁴ /s ⁴)	α (1/m)
0	1.0E+05	3.0E-07	1.0E+05	3.0E-07	1.0E+05	3.0E-07
1	1.0E+03	7.7E-07	1.0E+03	6.0E-07	1.0E+03	6.0E-07
2	1.2E+02	3.0E-06	4.2E-01	5.0E-05	5.5E-01	5.0E-05
3	1.3E+01	8.5E-06	3.0E+01	1.0E-06	1.0E-02	1.0E-06
4	5.5E+00	2.0E-05	4.0E+01	1.0E-06	0.2E-02	1.0E-06
5	3.5E-01	6.0E-05	1.0E-01	0.8E-06	0.2E-02	0.8E-06
6	1.5E-02	1.2E-04	6.0E-04	9.7E-05	0.1E-02	9.7E-05
7	1.5E-04	2.0E-04	9.0E-04	2.5E-07	1.0E-03	2.5E-07
8	2.2E-05	5.0E-04	7.1E-08	0.8E-03	8.1E-09	0.6E-03
9	2.1E-06	1.2E-03	6.0E-10	4.1E-02	1.0E-14	2.1E-02
10	2.0E-07	3.0E-03	6.0E-10	0.4E-02	1.0E-14	0.4E-02
11	3.0E-08	6.5E-02	6.0E-10	1.1E-02	1.0E-14	1.1E-02
12	1.0E-10	1.9E-02	1.0E-10	4.8E-02	1.0E-14	1.8E-02
13	1.4E-10	6.0E+01	1.0E-10	1.2E-01	1.0E-14	1.2E-01
14	1.5E-10	1.7E+03	1.0E-13	2.1E-01	1.0E-14	2.1E-01
15	1.5E-13	5.7E+03	1.0E-10	8.0E-01	1.0E-14	8.0E-01
Altitude of vehicle	$x_3 = x'_3 = 355$ m					
N_f	e.g., 300 [E/(cyc/m)] ² for the speed of 10m/sec at the sampling rate of 3 seconds) (1E/ $\sqrt{\text{Hz}}$ per data track)					

Table A.1: Values of model parameters for the gravitational gradients

APPENDIX B

A RECTANGULAR PRISM AND ITS PARTIAL DERIVATIVES WITH RESPECT TO ORIGIN AND ORIENTATION

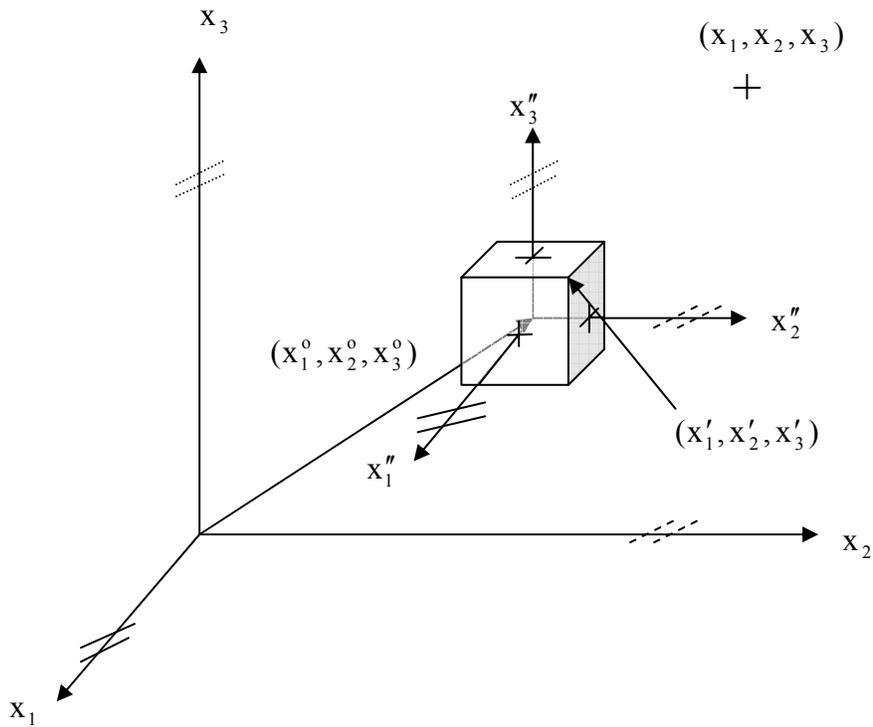


Figure B.1: A rectangular prism represents an anomalous mass measured at point (x_1, x_2, x_3)

B.1 A rectangular prism

In Figure B.1, suppose a rectangular prism is held fix with \mathbf{x}'' -triad frame parallel to a local coordinate system: $(1,2,3) = (N,E,D)$, called \mathbf{x} -triad frame. The prism has the dimension of $2\Delta x_1'' \times 2\Delta x_2'' \times 2\Delta x_3''$ and its center is at (x_1^0, x_2^0, x_3^0) (see also Figure B.2).

Under a plane assumption, analytic formulae of gravitational gradient components with cyclic index $(i, j) = (1, 2, 3)$ due to a rectangular prism are given by [Forsberg (1984), Jekeli et al. (2003), and Zhdanov et al. (2004)]:

$$\Gamma_{ii}(\mathbf{x}) = k\Delta\rho \tan^{-1} \left[\frac{(x_j - x'_j)(x_k - x'_k)}{(x_i - x'_i)r} \right] \Bigg|_{x'_{i,1}}^{x'_{i,2}} \Bigg|_{x'_{j,1}}^{x'_{j,2}} \Bigg|_{x'_{k,1}}^{x'_{k,2}} \quad (\text{B.1})$$

and

$$\Gamma_{ij}(\mathbf{x}) = -k\Delta\rho \ln[(x_k - x'_k) + r] \Bigg|_{x'_{i,1}}^{x'_{i,2}} \Bigg|_{x'_{j,1}}^{x'_{j,2}} \Bigg|_{x'_{k,1}}^{x'_{k,2}} \quad \text{with } (i, j, k) = (1, 2, 3) \quad (\text{B.2})$$

where

(x'_1, x'_2, x'_3) = a point on the prism with respect to \mathbf{x} -triad frame

k = Newtonian gravitational constant

$\Delta\rho$ = density contrast of the prism

$$r = \sqrt{(x_i - x'_i)^2 + (x_j - x'_j)^2 + (x_k - x'_k)^2}$$

B.2 Origin and orientation

Suppose the rectangular prism have an orientation, α , about x''_3 -axis, positive in counter-clockwise direction with respect to \mathbf{x}'' -triad frame, at view of end (see Figure B.2). The transformation of six gradients of tensor from \mathbf{x}'' -frame to \mathbf{x} -frame can be viewed such that \mathbf{x}'' -frame is oriented, parallel to \mathbf{x} -frame by α at origin (x''_1, x''_2, x''_3) . The transformation of the gravitational gradient tensor from $\mathbf{\Gamma}''$ to $\mathbf{\Gamma}$ is

$$\mathbf{\Gamma} = \mathbf{R}(\alpha)\mathbf{\Gamma}''\mathbf{R}(\alpha)^T \quad (\text{B.3})$$

with rotation matrix

$$\mathbf{R}(\alpha) = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) & 0 \\ -\sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (\text{B.4})$$

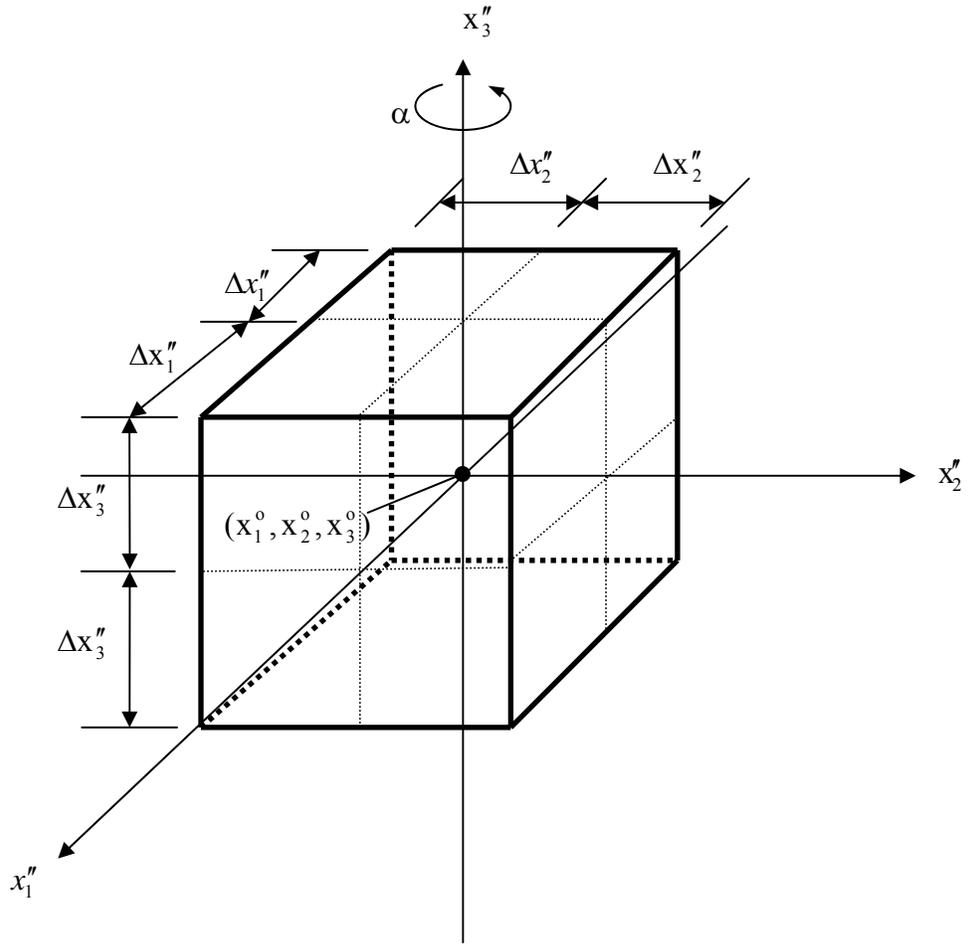


Figure B.2: A rectangular prism is held fixed with respect to \mathbf{x}'' -triad frame

For this work, origin (x_1^0, x_2^0, x_3^0) and orientation of the prism are treated as unknown parameters. The partial derivatives of gradient components with respect to the parameters can be derived as follows. We can write (B.4) as a linear form associated with the unknown parameters, i.e., $\boldsymbol{\theta} = [x_1^0, x_2^0, x_3^0, \alpha]^T$, as follows:

$$\begin{bmatrix} \Gamma_{33}(\mathbf{x}; \boldsymbol{\theta}) \\ \Gamma_{22}(\mathbf{x}; \boldsymbol{\theta}) \\ \Gamma_{11}(\mathbf{x}; \boldsymbol{\theta}) \\ \Gamma_{12}(\mathbf{x}; \boldsymbol{\theta}) \\ \Gamma_{13}(\mathbf{x}; \boldsymbol{\theta}) \\ \Gamma_{23}(\mathbf{x}; \boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_\alpha^2 & s_\alpha^2 & -2c_\alpha s_\alpha & 0 & 0 \\ 0 & s_\alpha^2 & c_\alpha^2 & +2c_\alpha s_\alpha & 0 & 0 \\ 0 & +c_\alpha s_\alpha & -c_\alpha s_\alpha & (c_\alpha^2 - s_\alpha^2) & 0 & 0 \\ 0 & 0 & 0 & 0 & c_\alpha & s_\alpha \\ 0 & 0 & 0 & 0 & -s_\alpha & c_\alpha \end{bmatrix} \begin{bmatrix} \Gamma_{33}''(\mathbf{x}; x_1^0, x_2^0, x_3^0) \\ \Gamma_{22}''(\mathbf{x}; x_1^0, x_2^0, x_3^0) \\ \Gamma_{11}''(\mathbf{x}; x_1^0, x_2^0, x_3^0) \\ \Gamma_{12}''(\mathbf{x}; x_1^0, x_2^0, x_3^0) \\ \Gamma_{13}''(\mathbf{x}; x_1^0, x_2^0, x_3^0) \\ \Gamma_{23}''(\mathbf{x}; x_1^0, x_2^0, x_3^0) \end{bmatrix} \quad (\text{B.5})$$

where

$$c_\alpha = \cos(\alpha) \text{ and } s_\alpha = \sin(\alpha). \quad (\text{B.6})$$

Let the unknown parameters $\boldsymbol{\theta} = \boldsymbol{\theta}_0 + \boldsymbol{\delta\theta}$, where $\boldsymbol{\theta}_0$ are approximation parameters and $\boldsymbol{\delta\theta}$ are perturbing terms. The gravitational gradient of component ij can be expanded to a Taylor's series with respect to $\boldsymbol{\theta}$ and given by

$$\Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta}) = \Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta}_0) + \left[\frac{\partial \Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}^T \boldsymbol{\delta\theta} \quad (\text{B.7})$$

where

$$\begin{aligned} \left[\frac{\partial \Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]_{\boldsymbol{\theta}=\boldsymbol{\theta}_0}^T \boldsymbol{\delta\theta} = & \frac{\partial \Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta})}{\partial x_1^0} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \delta x_1^0 + \frac{\partial \Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta})}{\partial x_2^0} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \delta x_2^0 + \frac{\partial \Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta})}{\partial x_3^0} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \delta x_3^0 \\ & + \frac{\partial \Gamma_{ij}(\mathbf{x}; \boldsymbol{\theta})}{\partial \alpha} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_0} \delta \alpha \end{aligned} \quad (\text{B.8})$$

For a convenience of derivations for each $\partial \Gamma_{ij} / \partial \boldsymbol{\theta}$, we define the following relations (alternatively, the derivatives can be considered in frequency domain, see Forsberg (1984) for more details):

(i) Diagonal element Γ_{ii}'' of gradient tensor in \mathbf{x}'' -frame

Let,

$$\mathbf{u}_{ii}^{1,m,n} = \tan^{-1}[\boldsymbol{\kappa}_{ii}^{1,m,n}] \quad (\text{B.9})$$

where

$$\boldsymbol{\kappa}_{ii}^{1,m,n} = \left[\frac{(x_j - x'_{j,m})(x_k - x'_{k,n})}{(x_i - x'_{i,l})r_{l,m,n}} \right] \text{ and} \quad (\text{B.10})$$

$$r_{l,m,n} = \sqrt{(x_i - x'_{i,l})^2 + (x_j - x'_{j,m})^2 + (x_k - x'_{k,n})^2} \quad (\text{B.11})$$

$$x'_{i,l} = (-1)^l \Delta x_i'' + x_i^0 \text{ (similar for } x'_{j,m} \text{ and } x'_{k,n}) \quad (\text{B.12})$$

Then,

$$\frac{\partial \Gamma_{ii}''(\mathbf{x}; x_1^0, x_2^0, x_3^0)}{\partial x^0} = k \Delta \rho \sum_{l=1}^2 \sum_{m=1}^2 \sum_{n=1}^2 \Omega^{l,m,n} \left[\frac{\partial \mathbf{u}_{ii}^{1,m,n}}{\partial x^0} \right] \quad (\text{B.13})$$

where

$$\Omega^{l,m,n} = (-1)^l (-1)^m (-1)^n \quad (\text{B.14})$$

$$\frac{\partial u_{ii}^{1,m,n}}{\partial x_i^o} = \left[\frac{1}{1 + (\kappa_{ii}^{1,m,n})^2} \right] \left[\frac{(x_j - x'_{j,m})(x_k - x'_{k,n})}{(x_i - x'_{i,l})r_{1,m,n}} \right] \left[\frac{1}{(x_i - x'_{i,l})} + \frac{(x_i - x'_{i,l})}{r_{1,m,n}^2} \right] \quad (\text{B.15})$$

$$\frac{\partial u_{ii}^{1,m,n}}{\partial x_j^o} = \left[\frac{1}{1 + (\kappa_{ii}^{1,m,n})^2} \right] \left[\frac{(x_k - x'_{k,n})}{(x_i - x'_{i,l})r_{1,m,n}} \right] \left[-1 + \frac{(x_j - x'_{j,m})^2}{r_{1,m,n}^2} \right] \quad (\text{B.16})$$

and

$$\frac{\partial u_{ii}^{1,m,n}}{\partial x_k^o} = \left[\frac{1}{1 + (\kappa_{ii}^{1,m,n})^2} \right] \left[\frac{(x_j - x'_{j,m})}{(x_i - x'_{i,l})r_{1,m,n}} \right] \left[-1 + \frac{(x_k - x'_{k,n})^2}{r_{1,m,n}^2} \right] \quad (\text{B.17})$$

(ii) Off-diagonal element Γ'_{ij} of gradient tensor in \mathbf{x}'' -frame

Let,

$$\bar{u}_{ij}^{1,m,n} = \ln[\bar{\kappa}_{ij}^{1,m,n}] \quad (\text{B.18})$$

where

$$\bar{\kappa}_{ij}^{1,m,n} = [(x_k - x'_{k,n}) + r_{1,m,n}] \quad (\text{B.19})$$

Then, we write

$$\frac{\partial \Gamma''_{ij}(\mathbf{x}; x_1^o, x_2^o, x_3^o)}{\partial x^o} = -k\Delta\rho \sum_{l=1}^2 \sum_{m=1}^2 \sum_{n=1}^2 \Omega^{l,m,n} \left[\frac{\partial \bar{u}_{ij}^{1,m,n}}{\partial x^o} \right] \quad (\text{B.20})$$

where

$$\frac{\partial \bar{u}_{ij}^{1,m,n}}{\partial x_i^o} = \left[\frac{1}{\bar{\kappa}_{ij}^{1,m,n}} \right] \left[-\frac{(x_i - x'_{i,l})}{r_{1,m,n}} \right] \quad (\text{B.21})$$

$$\frac{\partial \bar{u}_{ij}^{1,m,n}}{\partial x_j^o} = \left[\frac{1}{\bar{\kappa}_{ij}^{1,m,n}} \right] \left[-\frac{(x_j - x'_{j,m})}{r_{1,m,n}} \right] \quad (\text{B.22})$$

and

$$\frac{\partial \bar{u}_{ij}^{1,m,n}}{\partial x_k^o} = \left[\frac{1}{\bar{\kappa}_{ij}^{1,m,n}} \right] \left[-1 - \frac{(x_k - x'_{k,n})}{r_{1,m,n}} \right] \quad (\text{B.23})$$

Then, each element $\partial \Gamma_{ij} / \partial \theta$ of (B.8) can be expressed as the following.

(1) For Γ_{33} ,

$$\Gamma_{33} = \Gamma''_{33} \quad (\text{B.24})$$

$$\frac{\partial \Gamma_{33}}{\partial x_i^o} = \frac{\partial \Gamma''_{33}}{\partial x_i^o} \quad \text{with } i = 1, 2, 3 \quad (\text{B.25})$$

$$\frac{\partial \Gamma_{33}}{\partial \alpha} = 0 \quad (\text{B.26})$$

(2) For Γ_{22} ,

$$\Gamma_{22} = \mathbf{c}_\alpha^2 \Gamma_{22}'' + \mathbf{s}_\alpha^2 \Gamma_{11}'' - 2\mathbf{c}_\alpha \mathbf{s}_\alpha \Gamma_{12}'' \quad (\text{B.27})$$

$$\frac{\partial \Gamma_{22}}{\partial \mathbf{x}_i^0} = \mathbf{c}_\alpha^2 \frac{\partial \Gamma_{22}''}{\partial \mathbf{x}_i^0} + \mathbf{s}_\alpha^2 \frac{\partial \Gamma_{11}''}{\partial \mathbf{x}_i^0} - 2\mathbf{c}_\alpha \mathbf{s}_\alpha \frac{\partial \Gamma_{12}''}{\partial \mathbf{x}_i^0} \quad (\text{see the equations of (B.13) and (B.20)}) \quad (\text{B.28})$$

$$\frac{\partial \Gamma_{22}}{\partial \alpha} = -2\mathbf{c}_\alpha \mathbf{s}_\alpha \Gamma_{22}'' + 2\mathbf{c}_\alpha \mathbf{s}_\alpha \Gamma_{11}'' + 2(\mathbf{s}_\alpha^2 - \mathbf{c}_\alpha^2) \Gamma_{12}'' \quad (\text{B.29})$$

(3) For Γ_{11} ,

$$\Gamma_{11} = \mathbf{s}_\alpha^2 \Gamma_{22}'' + \mathbf{c}_\alpha^2 \Gamma_{11}'' + 2\mathbf{c}_\alpha \mathbf{s}_\alpha \Gamma_{12}'' \quad (\text{B.30})$$

$$\frac{\partial \Gamma_{11}}{\partial \mathbf{x}_i^0} = \mathbf{s}_\alpha^2 \frac{\partial \Gamma_{22}''}{\partial \mathbf{x}_i^0} + \mathbf{c}_\alpha^2 \frac{\partial \Gamma_{11}''}{\partial \mathbf{x}_i^0} + 2\mathbf{c}_\alpha \mathbf{s}_\alpha \frac{\partial \Gamma_{12}''}{\partial \mathbf{x}_i^0} \quad (\text{see the equations of (B.13) and (B.20)}) \quad (\text{B.31})$$

$$\frac{\partial \Gamma_{11}}{\partial \alpha} = 2\mathbf{s}_\alpha \mathbf{c}_\alpha (\Gamma_{22}'' - \Gamma_{11}'') + 2(\mathbf{c}_\alpha^2 - \mathbf{s}_\alpha^2) \Gamma_{12}'' \quad (\text{B.32})$$

(4) For Γ_{12} ,

$$\Gamma_{12} = \mathbf{c}_\alpha \mathbf{s}_\alpha \Gamma_{22}'' - \mathbf{c}_\alpha \mathbf{s}_\alpha \Gamma_{11}'' + (\mathbf{c}_\alpha^2 - \mathbf{s}_\alpha^2) \Gamma_{12}'' \quad (\text{B.33})$$

$$\frac{\partial \Gamma_{12}}{\partial \mathbf{x}_i^0} = \mathbf{c}_\alpha \mathbf{s}_\alpha \frac{\partial \Gamma_{22}''}{\partial \mathbf{x}_i^0} - \mathbf{c}_\alpha \mathbf{s}_\alpha \frac{\partial \Gamma_{11}''}{\partial \mathbf{x}_i^0} + (\mathbf{c}_\alpha^2 - \mathbf{s}_\alpha^2) \frac{\partial \Gamma_{12}''}{\partial \mathbf{x}_i^0} \quad (\text{see the equations of (B.13) and (B.20)}) \quad (\text{B.34})$$

$$\frac{\partial \Gamma_{12}}{\partial \alpha} = (\mathbf{c}_\alpha^2 - \mathbf{s}_\alpha^2) \Gamma_{22}'' - (\mathbf{c}_\alpha^2 - \mathbf{s}_\alpha^2) \Gamma_{11}'' + 4\mathbf{c}_\alpha \mathbf{s}_\alpha \Gamma_{12}'' \quad (\text{B.35})$$

(5) For Γ_{13} ,

$$\Gamma_{13} = \mathbf{c}_\alpha \Gamma_{13}'' + \mathbf{s}_\alpha \Gamma_{23}'' \quad (\text{B.36})$$

$$\frac{\partial \Gamma_{13}}{\partial \mathbf{x}_i^0} = \mathbf{c}_\alpha \frac{\partial \Gamma_{13}''}{\partial \mathbf{x}_i^0} + \mathbf{s}_\alpha \frac{\partial \Gamma_{23}''}{\partial \mathbf{x}_i^0} \quad (\text{see the equations of (B.13) and (B.20)}) \quad (\text{B.37})$$

$$\frac{\partial \Gamma_{13}}{\partial \alpha} = -\mathbf{s}_\alpha \Gamma_{13}'' + \mathbf{c}_\alpha \Gamma_{23}'' \quad (\text{B.38})$$

(6) For Γ_{23} ,

$$\Gamma_{23} = -\mathbf{s}_\alpha \Gamma_{13}'' + \mathbf{c}_\alpha \Gamma_{23}'' \quad (\text{B.39})$$

$$\frac{\partial \Gamma_{23}}{\partial x_i^0} = -s_\alpha \frac{\partial \Gamma_{13}''}{\partial x_i^0} + c_\alpha \frac{\partial \Gamma_{23}''}{\partial x_i^0} \quad (\text{see the equations of (B.12) and (B.20)}) \quad (\text{B.40})$$

$$\frac{\partial \Gamma_{23}}{\partial \alpha} = -c_\alpha \Gamma_{13}'' - s_\alpha \Gamma_{23}'' \quad (\text{B.41})$$

B.3 The partial derivatives of a rectangular prism in the frequency domain

As can be seen already, the partial derivatives of a rectangular prism are very complicated in the space domain. Surprisingly, they are simple in the frequency domain. We simply take the Fourier transform, \mathfrak{F} , of the gradient tensors due to a rectangular with respect to (x_1, x_2) . We have the Fourier transform pair:

$$\Gamma_{ij}(x_1, x_2; x_3) \Leftrightarrow \mathfrak{F}(\Gamma_{ij}(x_1, x_2; x_3)) = \tilde{\Gamma}_{ij}(f_1, f_2), \quad (\text{B.42})$$

$$(x_1, x_2) \Leftrightarrow (f_1, f_2) \quad (\text{B.43})$$

with

$$\mathfrak{F}(\cdot) = \int_{-\infty-\infty}^{+\infty+\infty} \int (\cdot) e^{-2\pi i(x_1 f_1 + x_2 f_2)} dx_1 dx_2 \quad (\text{B.44})$$

Thus, the partial derivatives with respect to x_m , for $m = (1, 2)$, can be readily obtained and given by

$$\frac{\partial \Gamma_{ij}}{\partial x_m} = \mathfrak{F}^{-1}[(2\pi i x_m) \tilde{\Gamma}_{ij}(f_1, f_2)] \quad (\text{B.45})$$

or

$$\frac{\partial \Gamma_{ij}}{\partial x'_m} = \mathfrak{F}^{-1}[(2\pi i x'_m)^* \tilde{\Gamma}_{ij}(f_1, f_2)] \quad (\text{B.46})$$

where \mathfrak{F}^{-1} is the inverse of Fourier transform operator

$$\mathfrak{F}^{-1}(\cdot) = \int_{-\infty-\infty}^{+\infty+\infty} \int (\cdot) e^{2\pi i(x_1 f_1 + x_2 f_2)} df_1 df_2 \quad (\text{B.47})$$

The partial derivatives with respect to x_3 is given by

$$\frac{\partial \Gamma_{ij}}{\partial x_3} = \frac{\partial \Gamma_{ij}}{\partial x'_3} = \mathfrak{F}^{-1}[(-2\pi f) \tilde{\Gamma}_{ij}(f_1, f_2)] \quad \text{with } f = \sqrt{f_1^2 + f_2^2} \quad (\text{B.48})$$

APPENDIX C

REARRANGMENT OF DATA TRACKS BY APPLYING ORTHOGONAL TRANSFORMATION

In this section, we show comprehensive details of rearrangement of gradient data on the basis of orthogonal transformations and permutations for various cases of data tracks and sensors.

C.1 Single data track and single sensor (SS)

For a convenience, we write the orthogonal matrix \mathbf{F} containing the N orthonormal basis vectors $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{N-1}$, defined in (6.22) through (6.24) in chapter 6:

$$\mathbf{F} = [\mathbf{v}_0 \quad \mathbf{v}_1 \quad \dots \quad \mathbf{v}_{N-1}] \quad (\text{C.1})$$

with

$$\mathbf{v}_p = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 \\ e^{i2\pi\Delta x f_p} \\ \vdots \\ e^{i2\pi\Delta x (N-1)f_p} \end{bmatrix} \quad (\text{C.2})$$

and

$$\mathbf{v}_{N-p} = \mathbf{v}_p^* \quad (\text{C.3})$$

where $f_p = \frac{p}{N\Delta x}$ for $p = 0, 1, 2, \dots, N-1$, Δx is sampling interval, and the asterisk denotes complex conjugate. Then,

$$\mathbf{F}^H \mathbf{F} = \mathbf{F} \mathbf{F}^H = \mathbf{I} \quad \text{and} \quad \mathbf{F}^H = \mathbf{F}^{-1} \quad (\text{C.4})$$

For a given $N \times N$ covariance matrix of gradient noise

$$\boldsymbol{\Sigma}_{SS} = \begin{bmatrix} r(0) & r(-1) & \dots & r(-(N-1)) \\ r(1) & r(0) & \dots & r(-N) \\ \vdots & \vdots & \ddots & \vdots \\ r(N-1) & r(N) & \dots & r(0) \end{bmatrix} \quad (\text{C.5})$$

or in a compact form

$$\mathbf{\Sigma}_{SS} = [r(i-j)]_{ij} \quad (\text{C.6})$$

where $r(i-j)$ is a correlation function for $i, j=0,1,\dots,N-1$. For zero-mean and wide-sense stationary process, it can be shown that the product of $\mathbf{F}^H \mathbf{\Sigma}_{SS} \mathbf{F}$ is a diagonal matrix:

$$\begin{aligned} \tilde{\mathbf{\Sigma}}_{SS} &= \mathbf{F}^H \mathbf{\Sigma}_{SS} \mathbf{F} = [\mathbf{v}_0 \quad \mathbf{v}_1 \quad \dots \quad \mathbf{v}_{N-1}]^H \mathbf{\Sigma}_{SS} \mathbf{F} \\ &= \frac{1}{\sqrt{N}} \begin{bmatrix} \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_0 \Delta x(k+0)} & \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_0 \Delta x(k+1)} & \dots & \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_0 \Delta x(k+(N-1))} \\ \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_1 \Delta x(k+0)} & \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_1 \Delta x(k+1)} & \dots & \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_1 \Delta x(k+(N-1))} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_{N-1} \Delta x(k+0)} & \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_{N-1} \Delta x(k+1)} & \dots & \sum_{k=0}^{N-1} r(k) e^{-i2\pi f_{N-1} \Delta x(k+(N-1))} \end{bmatrix} \mathbf{F} \\ &= \begin{bmatrix} \left(\sum_{k=0}^{N-1} r(k) e^{-i2\pi f_0 \Delta x k} \right) \mathbf{v}_0^H \\ \left(\sum_{k=0}^{N-1} r(k) e^{-i2\pi f_1 \Delta x k} \right) \mathbf{v}_1^H \\ \vdots \\ \left(\sum_{k=0}^{N-1} r(k) e^{-i2\pi f_{N-1} \Delta x k} \right) \mathbf{v}_{N-1}^H \end{bmatrix} [\mathbf{v}_0 \quad \mathbf{v}_1 \quad \dots \quad \mathbf{v}_{N-1}] \\ &= \begin{bmatrix} \sum_{k=0}^{N-1} r(k) e^{-i2\pi k \Delta x f_0} & & & 0 \\ & \sum_{k=0}^{N-1} r(k) e^{-i2\pi k \Delta x f_1} & & \\ & & \ddots & \\ 0 & & & \sum_{k=0}^{N-1} r(k) e^{-i2\pi k \Delta x f_{N-1}} \end{bmatrix} \end{aligned} \quad (\text{C.7})$$

C.2 Single data track and multiple sensors (SM)

If we have K multiple sensors, the case of single track can be easily extended as the following. Given a large data vector

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{K-1} \end{bmatrix} = \begin{bmatrix} y_0(x_0) \\ y_0(x_1) \\ \vdots \\ y_0(x_{N-1}) \\ \hline y_1(x_0) \\ y_1(x_1) \\ \vdots \\ y_1(x_{N-1}) \\ \hline \vdots \\ \hline y_{K-1}(x_0) \\ y_{K-1}(x_1) \\ \vdots \\ y_{K-1}(x_{N-1}) \end{bmatrix} \begin{array}{l} \text{sensor 0} \\ \\ \\ \text{sensor 1} \\ \\ \\ \text{sensor } K-1 \end{array} \quad (\text{C.8})$$

with $MN \times MN$ cross-covariance matrix

$$\boldsymbol{\Sigma}_{SM} = [\boldsymbol{\Sigma}_{ij}] \text{ with } (i, j) = 0, 1, \dots, K-1 \quad (\text{C.9})$$

where the $[i, j]$ block is the cross-covariance matrix between sensors i and j . Each block matrix has dimension $N \times N$ and its elements are a cross-covariance (or cross-correlation) function r_{ij} of sensors i and j

$$\boldsymbol{\Sigma}_{ij} = [r_{ij}(m-n)]_{m,n} \text{ for } (m, n) = 0, 1, \dots, N-1 \quad (\text{C.10})$$

Let the $KN \times KN$ transformation matrix be defined as

$$\mathbf{F}_{SM} = \text{diag}(\mathbf{F}_0, \mathbf{F}_1, \dots, \mathbf{F}_i, \dots, \mathbf{F}_{K-1}) \quad (\text{C.11})$$

where $\mathbf{F}_i = \mathbf{F}$ (defined in (C.1)) for $i = 0, 1, \dots, N-1$, and the $KN \times KN$ permutation matrix \mathbf{Z} (\mathbf{Z} also is an orthogonal matrix) is constructed from the following algorithm (in MATLAB format)

```

M' = K
N' = N
k = 1
for i = 0:N'-1
    for j = 0:M'-1
        Z(:,k) = I(:,i+(j-1)*N');
        k = 1+k;
    end
end
end

```

(C.12)

Since the columns of Z form an orthonormal set (with respect to the usual inner product), Z is a nonsingular and orthonormal matrix (Harville, 1997). The following expressions, from (C.13) to (C.17), give an example of permuting a data vector together with its covariance matrix. Let $K = 3$ and $N = 2$ and a data vector be given by

$$\mathbf{y} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix} \begin{array}{l} \text{sensor 1} \\ \text{-----} \\ \text{sensor 2} \\ \text{-----} \\ \text{sensor 3} \end{array} \quad (C.13)$$

with the covariance matrix

$$\mathbf{C} = \begin{bmatrix} r_{11} & 0 & r_{12} & 0 & r_{13} & 0 \\ 0 & r_{11} & 0 & r_{12} & 0 & r_{13} \\ r_{21} & 0 & r_{22} & 0 & r_{23} & 0 \\ 0 & r_{21} & 0 & r_{22} & 0 & r_{23} \\ r_{31} & 0 & r_{32} & 0 & r_{33} & 0 \\ 0 & r_{31} & 0 & r_{32} & 0 & r_{33} \end{bmatrix}, \quad (C.14)$$

The permutation matrix constructed using the algorithm of (C.12) is

$$\mathbf{Z} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (C.15)$$

Then,

$$\mathbf{Z}^T \mathbf{y} = [y_{11} \quad y_{21} \quad y_{31} \quad | \quad y_{12} \quad y_{22} \quad y_{32}]^T \quad (\text{C.16})$$

with

$$\mathbf{Z}^T \mathbf{C} \mathbf{Z} = \begin{bmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & \mathbf{r}_{13} & & & \\ \mathbf{r}_{21} & \mathbf{r}_{22} & \mathbf{r}_{23} & & \mathbf{O} & \\ \mathbf{r}_{31} & \mathbf{r}_{32} & \mathbf{r}_{33} & & & \\ \hline & & & \mathbf{r}_{11} & \mathbf{r}_{12} & \mathbf{r}_{13} \\ & \mathbf{O} & & \mathbf{r}_{21} & \mathbf{r}_{22} & \mathbf{r}_{23} \\ & & & \mathbf{r}_{31} & \mathbf{r}_{32} & \mathbf{r}_{33} \end{bmatrix} \quad (\text{C.17})$$

Since

$$(\mathbf{Z} \mathbf{F}_{\text{SM}})(\mathbf{Z} \mathbf{F}_{\text{SM}})^H = (\mathbf{Z} \mathbf{F}_{\text{SM}})^H (\mathbf{Z} \mathbf{F}_{\text{SM}}) = \mathbf{I}_{\text{KN}} \quad \text{and} \quad (\mathbf{Z} \mathbf{F}_{\text{SM}})^H = (\mathbf{Z} \mathbf{F}_{\text{SM}})^{-1} \quad (\text{C.18})$$

the following steps show that the product of $\mathbf{Z}^H \mathbf{F}_{\text{SM}}^H \boldsymbol{\Sigma}_{\text{SM}} \mathbf{F}_{\text{SM}} \mathbf{Z}$ is a block diagonal matrix

Step I:

$$\mathbf{F}_{\text{SM}}^H \boldsymbol{\Sigma}_{\text{SM}} \mathbf{F}_{\text{SM}} = [\mathbf{F}^H \boldsymbol{\Sigma}_{ij} \mathbf{F}]_{ij} \quad \text{for } (i, j) = 0, 1, \dots, N-1 \quad (\text{C.19})$$

where the $[i, j]$ block $\mathbf{F}^H \boldsymbol{\Sigma}_{ij} \mathbf{F}$ renders a $N \times N$ diagonal matrix

$$\mathbf{F}^H \boldsymbol{\Sigma}_{ij} \mathbf{F} = \text{diag} \left(\sum_{k=0}^{N-1} r_{ij}(k) e^{-i2\pi k \Delta x f_0}, \sum_{k=0}^{N-1} r_{ij}(k) e^{-i2\pi k \Delta x f_1}, \dots, \sum_{k=0}^{N-1} r_{ij}(k) e^{-i2\pi k \Delta x f_{N-1}} \right) \quad (\text{C.20})$$

Step II:

After rearrangement, the transformed covariance matrix $\tilde{\boldsymbol{\Sigma}}_{\text{SM}}$ of $\boldsymbol{\Sigma}_{\text{SM}}$ is a block diagonal matrix. It can be done by applying the $\text{KN} \times \text{KN}$ permutation matrix \mathbf{Z} , constructed by using (C.12) with $M' = \text{K}$ and $N' = \text{N}$:

$$\tilde{\boldsymbol{\Sigma}}_{\text{SM}} = \mathbf{Z}^H \mathbf{F}_{\text{SM}}^H \boldsymbol{\Sigma}_{\text{SM}} \mathbf{F}_{\text{SM}} \mathbf{Z} = \text{diag} \left(\tilde{\boldsymbol{\Sigma}}_{\text{SM}}(f_0), \tilde{\boldsymbol{\Sigma}}_{\text{SM}}(f_1), \dots, \tilde{\boldsymbol{\Sigma}}_{\text{SM}}(f_i), \dots, \tilde{\boldsymbol{\Sigma}}_{\text{SM}}(f_{N-1}) \right) \quad (\text{C.21})$$

where the i^{th} block $\tilde{\boldsymbol{\Sigma}}_{\text{SM}}(f_i)$ is

$$\tilde{\Sigma}_{SM}(\mathbf{f}_i) = \begin{bmatrix} \sum_{k=0}^{N-1} \mathbf{r}_{00}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} & \sum_{k=0}^{N-1} \mathbf{r}_{01}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} & \cdots & \sum_{k=0}^{N-1} \mathbf{r}_{0(K-1)}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} \\ \sum_{k=0}^{N-1} \mathbf{r}_{10}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} & \sum_{k=0}^{N-1} \mathbf{r}_{11}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} & \cdots & \sum_{k=0}^{N-1} \mathbf{r}_{1(K-1)}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} \mathbf{r}_{(K-1)0}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} & \sum_{k=0}^{N-1} \mathbf{r}_{(K-1)1}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} & \cdots & \sum_{k=0}^{N-1} \mathbf{r}_{(K-1)(K-1)}(\mathbf{k}) e^{-i2\pi\mathbf{k}\Delta\mathbf{x}\mathbf{f}_i} \end{bmatrix} \quad (\text{C.22})$$

Thus, we also have

$$\tilde{\mathbf{y}}_{SM} = \mathbf{Z}_{SM}^H \mathbf{F}_{SM}^H \mathbf{y} = \begin{bmatrix} \tilde{\mathbf{y}}(\mathbf{f}_0) \\ \tilde{\mathbf{y}}(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{y}}(\mathbf{f}_{N-1}) \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{y}}_0(\mathbf{f}_0) \\ \tilde{\mathbf{y}}_1(\mathbf{f}_0) \\ \vdots \\ \tilde{\mathbf{y}}_{K-1}(\mathbf{f}_0) \\ \tilde{\mathbf{y}}_0(\mathbf{f}_1) \\ \tilde{\mathbf{y}}_1(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{y}}_{K-1}(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{y}}_0(\mathbf{f}_{N-1}) \\ \tilde{\mathbf{y}}_1(\mathbf{f}_{N-1}) \\ \vdots \\ \tilde{\mathbf{y}}_{K-1}(\mathbf{f}_{N-1}) \end{bmatrix} \quad (\text{C.23})$$

C.3 Multiple data tracks and single sensor (MS)

For a convenience, we write the data vector of (6.58):

$$\mathbf{y} = \begin{bmatrix} y(0,0) \\ y(1,0) \\ \vdots \\ y(N-1,0) \\ \hline y(0,1) \\ y(1,1) \\ \vdots \\ y(N-1,1) \\ \hline \vdots \\ \hline y(0,M-1) \\ y(1,M-1) \\ \vdots \\ y(N-1,M-1) \end{bmatrix} \begin{array}{l} \text{track 0} \\ \dots \\ \text{track 1} \\ \dots \\ \text{track } M-1 \end{array} \quad (\text{C.24})$$

and the $M \times N$ covariance matrix, Σ_{MS} , of (6.59), where M is the number of data tracks and N is the number of data samples per track,

$$\Sigma_{MS} = [\Sigma_{pq}] \text{ for } p, q = 0, 1, \dots, M-1 \quad (\text{C.25a})$$

The Σ_{pq} is a $N \times N$ block cross-covariance matrix of tracks i and j

$$\Sigma_{pq} = [r_{pq}(i-j, p-q)]_{ij} \text{ for } i, j = 0, 1, \dots, N-1 \quad (\text{C.25b})$$

where $r(i-j, p-q)$ is a correlation function. We also assume sampling interval is Δx and track spacing is Δy .

Given $MN \times MN$ orthogonal matrices,

$$\mathbf{F}_{MS} = \text{diag}(\mathbf{F}_0, \mathbf{F}_1, \dots, \mathbf{F}_i, \dots, \mathbf{F}_{M-1}) \quad (\text{C.26})$$

and

$$\mathbf{F}'_{MS} = \text{diag}(\mathbf{F}'_0, \mathbf{F}'_1, \dots, \mathbf{F}'_i, \dots, \mathbf{F}'_{N-1}) \quad (\text{C.27})$$

where the i^{th} block matrix, \mathbf{F}_i is defined as (C.1) and has dimension $N \times N$. The i^{th} block matrix \mathbf{F}'_i also is the same as (C.1) with the replacement of Δy and M and has dimension $M \times M$. Rearranging rows and columns of the product of $\mathbf{F}_{\text{MS}}^H \boldsymbol{\Sigma}_{\text{MS}} \mathbf{F}_{\text{MS}}$ using the $MN \times MN$ permutation matrix \mathbf{Z} , whose columns are obtained by permuting (rearranging) the columns of the $MN \times MN$ identity matrix \mathbf{I}_{MN} based on (C.12) with $M' = M$ and $N' = N$.

Since

$$(\mathbf{F}_{\text{MS}} \mathbf{Z} \mathbf{F}'_{\text{MS}})(\mathbf{F}_{\text{MS}} \mathbf{Z} \mathbf{F}'_{\text{MS}})^H = (\mathbf{F}_{\text{MS}} \mathbf{Z} \mathbf{F}'_{\text{MS}})^H (\mathbf{F}_{\text{MS}} \mathbf{Z} \mathbf{F}'_{\text{MS}}) = \mathbf{I}_{\text{MN}} \quad (\text{C.28})$$

and

$$(\mathbf{F}_{\text{MS}} \mathbf{Z} \mathbf{F}'_{\text{MS}})^H = (\mathbf{F}_{\text{MS}} \mathbf{Z} \mathbf{F}'_{\text{MS}})^{-1} \quad (\text{C.29})$$

it can be shown that the product of $\mathbf{F}'_{\text{MS}}{}^H \mathbf{Z}^H \mathbf{F}_{\text{MS}}^H \boldsymbol{\Sigma}_{\text{MS}} \mathbf{F}_{\text{MS}} \mathbf{Z} \mathbf{F}'_{\text{MS}}$ is a block diagonal matrix as follows:

Step I

$$\mathbf{F}_{\text{MS}}^H \boldsymbol{\Sigma}_{\text{MS}} \mathbf{F}_{\text{MS}} = [\mathbf{F}^H \boldsymbol{\Sigma}_{pq} \mathbf{F}]_{pq} \quad \text{for } p, q = 0, 1, \dots, M-1 \quad (\text{C.30})$$

where

$$\mathbf{F}^H \boldsymbol{\Sigma}_{pq} \mathbf{F} = \text{diag} \left(\sum_{k=0}^{N-1} r_{pq}(k, p-q) e^{-i2\pi k \Delta x f_0}, \sum_{k=0}^{N-1} r_{pq}(k, p-q) e^{-i2\pi k \Delta x f_1}, \dots, \sum_{k=0}^{N-1} r_{pq}(k, p-q) e^{-i2\pi k \Delta x f_{N-1}} \right) \quad (\text{C.31})$$

Step II

$$\mathbf{Z}^H \mathbf{F}_{\text{MS}}^H \boldsymbol{\Sigma}_{\text{MS}} \mathbf{F}_{\text{MS}} \mathbf{Z} = \text{diag} \left(\tilde{\boldsymbol{\Sigma}}_{\text{MS}}(f_0), \tilde{\boldsymbol{\Sigma}}_{\text{MS}}(f_1), \dots, \tilde{\boldsymbol{\Sigma}}_{\text{MS}}(f_i), \dots, \tilde{\boldsymbol{\Sigma}}_{\text{MS}}(f_{N-1}) \right) \quad \text{for } i = 0, 1, \dots, N-1 \quad (\text{C.32})$$

where

$$\tilde{\boldsymbol{\Sigma}}_{\text{MS}}(f_i) = \begin{bmatrix} \sum_{k=0}^{N-1} r_{00}(k, 0) e^{-i2\pi k \Delta x f_i} & \sum_{k=0}^{N-1} r_{01}(k, -1) e^{-i2\pi k \Delta x f_i} & \dots & \sum_{k=0}^{N-1} r_{0, M-1}(k, -M+1) e^{-i2\pi k \Delta x f_i} \\ \sum_{k=0}^{N-1} r_{10}(k, 1) e^{-i2\pi k \Delta x f_i} & \sum_{k=0}^{N-1} r_{11}(k, 0) e^{-i2\pi k \Delta x f_i} & \dots & \sum_{k=0}^{N-1} r_{1, M-1}(k, -M) e^{-i2\pi k \Delta x f_i} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{N-1} r_{M-1, 0}(k, M-1) e^{-i2\pi k \Delta x f_i} & \sum_{k=0}^{N-1} r_{M-1, 1}(k, M-2) e^{-i2\pi k \Delta x f_i} & \dots & \sum_{k=0}^{N-1} r_{M-1, M-1}(k, 0) e^{-i2\pi k \Delta x f_i} \end{bmatrix} \quad (\text{C.33})$$

Step III

$$\begin{aligned}\tilde{\Sigma}_{MS} &= \mathbf{F}'_M \mathbf{Z}^H \mathbf{F}_M^H \Sigma_{MS} \mathbf{F}_M \mathbf{Z} \mathbf{F}'_M \\ &= \text{diag}(\mathbf{F}'^H \tilde{\Sigma}_{MS}(f_0) \mathbf{F}', \mathbf{F}'^H \tilde{\Sigma}_{MS}(f_1) \mathbf{F}', \dots, \mathbf{F}'^H \tilde{\Sigma}_{MS}(f_i) \mathbf{F}', \dots, \mathbf{F}'^H \tilde{\Sigma}_{MS}(f_{N-1}) \mathbf{F}')\end{aligned}\quad (\text{C.34})$$

where

$$\begin{aligned}\mathbf{F}'^H \tilde{\Sigma}_{MS}(f_i) \mathbf{F}' &= \begin{bmatrix} \left(\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(k, k') e^{-i2\pi(f_i \Delta x k + f'_0 \Delta y k')} \right) \mathbf{v}'_0{}^H \\ \left(\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(k, k') e^{-i2\pi(f_i \Delta x k + f'_1 \Delta y k')} \right) \mathbf{v}'_1{}^H \\ \vdots \\ \left(\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(k, k') e^{-i2\pi(f_i \Delta x k + f'_{M-1} \Delta y k')} \right) \mathbf{v}'_{M-1}{}^H \end{bmatrix} \begin{bmatrix} \mathbf{v}'_0 & \mathbf{v}'_1 & \dots & \mathbf{v}'_{M-1} \end{bmatrix} \\ &= \text{diag} \left(\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(k, k') e^{-i2\pi(f_i \Delta x k + f'_0 \Delta y k')}, \sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(k, k') e^{-i2\pi(f_i \Delta x k + f'_1 \Delta y k')}, \dots, \sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r(k, k') e^{-i2\pi(f_i \Delta x k + f'_{M-1} \Delta y k')} \right)\end{aligned}\quad (\text{C.35})$$

We have (see also (6.60) in chapter 6)

$$\tilde{\mathbf{y}}_{MS} = \begin{bmatrix} \tilde{\mathbf{y}}_{MS}(f_0) \\ \tilde{\mathbf{y}}_{MS}(f_1) \\ \vdots \\ \tilde{\mathbf{y}}_{MS}(f_i) \\ \vdots \\ \tilde{\mathbf{y}}_{MS}(f_{N-1}) \end{bmatrix} \quad \text{with } \tilde{\mathbf{y}}_{MS}(f_i) = \begin{bmatrix} \tilde{y}(f_i, f'_0) \\ \tilde{y}(f_i, f'_1) \\ \vdots \\ \tilde{y}(f_i, f'_{M-1}) \end{bmatrix} \quad \text{for } i = 0, 1, \dots, N-1 \quad (\text{C.36})$$

C.4 Multiple data tracks and multiple sensors (MM)

For the case of multiple tracks, we assume we have a large number of tracks, say M , and N data samples per track and track spacing is Δy . For a convenience of derivations, the $KMN \times 1$ data vector of (6.72) in chapter 6 is written

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_i \\ \vdots \\ \mathbf{y}_{K-1} \end{bmatrix} \text{ with } \mathbf{y}_i = \begin{bmatrix} y_i(0,0) \\ y_i(1,0) \\ \vdots \\ y_i(N-1,0) \\ \hline y_i(0,1) \\ y_i(1,1) \\ \vdots \\ y_i(N-1,1) \\ \hline \vdots \\ \hline y_i(0,M-1) \\ y_i(1,M-1) \\ \vdots \\ y_i(N-1,M-1) \end{bmatrix} \text{ for } i = 0,1,\dots,K-1 \quad (\text{C.37})$$

and the $KM \times KM$ cross-covariance matrix of (6.73) is

$$\boldsymbol{\Sigma}_{MM} = \begin{bmatrix} \boldsymbol{\Sigma}_{00} & \boldsymbol{\Sigma}_{01} & \cdots & \boldsymbol{\Sigma}_{0j} & \cdots & \boldsymbol{\Sigma}_{0(K-1)} \\ \boldsymbol{\Sigma}_{10} & \boldsymbol{\Sigma}_{11} & \cdots & \boldsymbol{\Sigma}_{1j} & \cdots & \boldsymbol{\Sigma}_{1(K-1)} \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \boldsymbol{\Sigma}_{i0} & \boldsymbol{\Sigma}_{i1} & \cdots & \boldsymbol{\Sigma}_{ij} & \cdots & \boldsymbol{\Sigma}_{i(K-1)} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}_{(K-1)0} & \boldsymbol{\Sigma}_{(K-1)1} & \cdots & \boldsymbol{\Sigma}_{(K-1)j} & \cdots & \boldsymbol{\Sigma}_{(K-1)(K-1)} \end{bmatrix} \quad (\text{C.38})$$

The $[i,j]$ block, $\boldsymbol{\Sigma}_{ij}$, of (6.74), having dimension $MN \times MN$ is

$$\boldsymbol{\Sigma}_{ij} = \begin{bmatrix} \boldsymbol{\Sigma}(0,0) & \boldsymbol{\Sigma}(0,1) & \cdots & \boldsymbol{\Sigma}(0,n) & \cdots & \boldsymbol{\Sigma}(0,M-1) \\ \boldsymbol{\Sigma}(1,0) & \boldsymbol{\Sigma}(1,1) & \cdots & \boldsymbol{\Sigma}(1,n) & \cdots & \boldsymbol{\Sigma}(1,M-1) \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ \boldsymbol{\Sigma}(m,0) & \boldsymbol{\Sigma}(m,1) & \cdots & \boldsymbol{\Sigma}(m,n) & \cdots & \boldsymbol{\Sigma}(m,M-1) \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \boldsymbol{\Sigma}(M-1,0) & \boldsymbol{\Sigma}(M-1,1) & \cdots & \boldsymbol{\Sigma}(M-1,n) & \cdots & \boldsymbol{\Sigma}(M-1,M-1) \end{bmatrix} \quad (\text{C.39})$$

The $[n,m]$ block, $\Sigma(m,n)$, is the $N \times N$ cross-covariance matrix between tracks m and n is

$$\Sigma(m,n) = [r_{mn}(i-j, m-n)]_{ij} \text{ for } (i,j) = 0,1,\dots,N-1 \quad (\text{C.40})$$

where r_{mn} is a cross-correlation (covariance) function between tracks m and n .

Let us define \mathbf{F}_{MM} -transformation matrix, having dimension $KMN \times KMN$ and block diagonal structure,

$$\mathbf{F}_{MM} = \begin{bmatrix} \mathbf{F}_x & & & \mathbf{O} \\ & \mathbf{F}_x & & \\ & & \ddots & \\ \mathbf{O} & & & \mathbf{F}_x \end{bmatrix} \quad (\text{C.41})$$

where the block \mathbf{F}_x has dimension $MN \times MN$ and contains the $N \times N$ orthonormal matrix, \mathbf{F} , of (C.1) as

$$\mathbf{F}_x = \begin{bmatrix} \mathbf{F} & & & \mathbf{O} \\ & \mathbf{F} & & \\ & & \ddots & \\ \mathbf{O} & & & \mathbf{F} \end{bmatrix} \quad (\text{C.42})$$

And the $KMN \times KMN$ transformation matrix \mathbf{F}'_{MM} is defined as

$$\mathbf{F}'_{MM} = \begin{bmatrix} \mathbf{F}_y & & & \mathbf{O} \\ & \mathbf{F}_y & & \\ & & \ddots & \\ \mathbf{O} & & & \mathbf{F}_y \end{bmatrix} \quad (\text{C.43})$$

where the block \mathbf{F}_y has dimension $MN \times MN$ and contains the $M \times M$ orthonormal matrix, \mathbf{F}' , of (C.1) as

$$\mathbf{F}_y = \begin{bmatrix} \mathbf{F}' & & & \mathbf{O} \\ & \mathbf{F}' & & \\ & & \ddots & \\ \mathbf{O} & & & \mathbf{F}' \end{bmatrix} \quad (\text{C.44})$$

where \mathbf{F}' is defined as (C.1) with the replacement of Δy and M .

To rearrange rows and columns of (C.37) and (C.38), the $KMN \times KMN$ permutation matrices \mathbf{Z}_{MM} is defined as:

$$\mathbf{Z}_{MM} = \text{diag}(\mathbf{Z}_0, \mathbf{Z}_1, \dots, \mathbf{Z}_i, \dots, \mathbf{Z}_{M-1}) \quad (\text{C.45})$$

where the $MN \times MN$ permutation matrix $\mathbf{Z}_i = \mathbf{Z}$, based on (C.12) with $M' = M$ and $N' = N$ for $i = 0, 1, \dots, M-1$. The $KMN \times KMN$ permutation matrix \mathbf{Z}'_{MM} is constructed by using (C.12) with $M' = K$ and $N' = MN$.

Since

$$(\mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM}) (\mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM})^H = (\mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM})^H (\mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM}) = \mathbf{I}_{KMN} \quad (\text{C.46})$$

and

$$(\mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM})^H = (\mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM})^{-1} \quad (\text{C.47})$$

it can be shown that the product of $\mathbf{Z}'_{MM} \mathbf{F}'_{MM} \mathbf{Z}_{MM} \mathbf{F}_{MM} \Sigma_{MM} \mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM}$ yields a block diagonal matrix, say $\tilde{\Sigma}_{MM}$, as the following:

Step I

$$\mathbf{F}_{MM}^H \Sigma_{MM} \mathbf{F}_{MM} = \begin{bmatrix} \mathbf{F}_x^H \Sigma_{00} \mathbf{F}_x & \mathbf{F}_x^H \Sigma_{01} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{0j} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{0(K-1)} \mathbf{F}_x \\ \mathbf{F}_x^H \Sigma_{10} \mathbf{F}_x & \mathbf{F}_x^H \Sigma_{11} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{1j} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{1(K-1)} \mathbf{F}_x \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{F}_x^H \Sigma_{i0} \mathbf{F}_x & \mathbf{F}_x^H \Sigma_{i1} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{ij} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{i(K-1)} \mathbf{F}_x \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \mathbf{F}_x^H \Sigma_{(K-1)0} \mathbf{F}_x & \mathbf{F}_x^H \Sigma_{(K-1)1} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{(K-1)j} \mathbf{F}_x & \cdots & \mathbf{F}_x^H \Sigma_{(K-1)(K-1)} \mathbf{F}_x \end{bmatrix} \quad (\text{C.48})$$

where the $[i, j]$ block

$$\mathbf{F}_x^H \Sigma_{ij} \mathbf{F}_x = \begin{bmatrix} \mathbf{F}^H \Sigma(0,0) \mathbf{F} & \mathbf{F}^H \Sigma(0,1) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(0,n) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(0, M-1) \mathbf{F} \\ \mathbf{F}^H \Sigma(1,0) \mathbf{F} & \mathbf{F}^H \Sigma(1,1) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(1,n) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(1, M-1) \mathbf{F} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathbf{F}^H \Sigma(m,0) \mathbf{F} & \mathbf{F}^H \Sigma(m,1) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(m,n) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(m, M-1) \mathbf{F} \\ \vdots & \vdots & & \vdots & \ddots & \vdots \\ \mathbf{F}^H \Sigma(M-1,0) \mathbf{F} & \mathbf{F}^H \Sigma(M-1,1) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(M-1,n) \mathbf{F} & \cdots & \mathbf{F}^H \Sigma(M-1, M-1) \mathbf{F} \end{bmatrix} \quad (\text{C.49})$$

With orthogonal property of \mathbf{F} , therefore, the block $\mathbf{F}^H \boldsymbol{\Sigma}(m,n) \mathbf{F}$ is a diagonal matrix.

Step II

Next, we rearrange the transformed observation vector, $\mathbf{F}_{MM}^H \mathbf{y}$, and its corresponding covariance using the $KMN \times KMN$ block-diagonal matrix \mathbf{Z}_{MM} . We have

$$\mathbf{Z}_{MM}^H \mathbf{F}_{MM}^H \boldsymbol{\Sigma}_{MM} \mathbf{F}_{MM} \mathbf{Z}_{MM} = \begin{bmatrix} \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{00} \mathbf{F}_x \mathbf{Z} & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{01} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{0j} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{0(K-1)} \mathbf{F}_x \mathbf{Z} \\ \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{10} \mathbf{F}_x \mathbf{Z} & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{11} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{1j} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{1(K-1)} \mathbf{F}_x \mathbf{Z} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{i0} \mathbf{F}_x \mathbf{Z} & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{i1} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{ij} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{i(K-1)} \mathbf{F}_x \mathbf{Z} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{(K-1)0} \mathbf{F}_x \mathbf{Z} & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{(K-1)1} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{(K-1)j} \mathbf{F}_x \mathbf{Z} & \cdots & \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{(K-1)(K-1)} \mathbf{F}_x \mathbf{Z} \end{bmatrix} \quad (\text{C.50})$$

As a result, the product of $\mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{ij} \mathbf{F}_x \mathbf{Z}$ yields a $MN \times MN$ block-diagonal matrix.

Step III

We now apply the $MN \times MN$ transformation matrix \mathbf{F}'_{MM} of (C.43) to the transformed observation vector, i.e. $\mathbf{F}'_{MM} \mathbf{Z}_{MM}^H \mathbf{F}_{MM}^H \mathbf{y}$, and its respective covariance matrix

$$\mathbf{F}'_{MM} \mathbf{Z}_{MM}^H \mathbf{F}_{MM}^H \boldsymbol{\Sigma}_{MM} \mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} = \left[\mathbf{F}_y^H \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{ij} \mathbf{F}_x \mathbf{Z} \mathbf{F}_y \right]_{ij} \quad \text{for } (i, j) = 0, 1, \dots, K-1 \quad (\text{C.51})$$

where the $[i, j]$ block $\mathbf{F}_y^H \mathbf{Z}^H \mathbf{F}_x^H \boldsymbol{\Sigma}_{ij} \mathbf{F}_x \mathbf{Z} \mathbf{F}_y$ is a $MN \times MN$ diagonal matrix.

Step IV

Next, rearranging rows and columns of vector $\mathbf{F}'_{MM} \mathbf{Z}_{MM}^H \mathbf{F}_{MM}^H \mathbf{y}$ and $\mathbf{F}'_{MM} \mathbf{Z}_{MM}^H \mathbf{F}_{MM}^H \boldsymbol{\Sigma}_{MM} \mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM}$ using \mathbf{Z}'_{MM} in order that, after arrangement, the transformed covariance matrix becomes a block diagonal matrix, say $\tilde{\boldsymbol{\Sigma}}_{MM}$, in which each block has $K \times K$ dimension. That is

$$\begin{aligned} \tilde{\boldsymbol{\Sigma}}_{MM} &= \mathbf{Z}'_{MM} \mathbf{F}'_{MM} \mathbf{Z}_{MM}^H \mathbf{F}_{MM}^H \boldsymbol{\Sigma}_{MM} \mathbf{F}_{MM} \mathbf{Z}_{MM} \mathbf{F}'_{MM} \mathbf{Z}'_{MM} \\ &= \text{diag}(\tilde{\boldsymbol{\Sigma}}_{MM}(f_0), \tilde{\boldsymbol{\Sigma}}_{MM}(f_1), \dots, \tilde{\boldsymbol{\Sigma}}_{MM}(f_i), \dots, \tilde{\boldsymbol{\Sigma}}_{MM}(f_{N-1})) \end{aligned} \quad (\text{C.52})$$

where the i^{th} block $\tilde{\Sigma}_{\text{MM}}(\mathbf{f}_i)$ having dimension $\text{KM} \times \text{KM}$ is

$$\tilde{\Sigma}_{\text{MM}}(\mathbf{f}_i) = \text{diag}(\tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_0), \tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_1), \dots, \tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_j), \dots, \tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_{M-1})) \quad (\text{C.53})$$

with the $\text{K} \times \text{K}$ block

$$\tilde{\Sigma}(\mathbf{f}_i, \mathbf{f}'_j) = \left[\sum_{k'=0}^{M-1} \sum_{k=0}^{N-1} r_{\text{mm}}(k, k') e^{-i2\pi(\Delta x k f_i + \Delta y k' f'_j)} \right]_{\text{mm}} \quad \text{for } (m, n) = 0, 1, \dots, \text{K} - 1 \quad (\text{C.54})$$

We also have

$$\tilde{\mathbf{y}}_{\text{MM}} = \mathbf{Z}'_{\text{MM}} \mathbf{F}'_{\text{MM}} \mathbf{Z}_{\text{MM}} \mathbf{F}_{\text{MM}} \mathbf{y} = \begin{bmatrix} \tilde{\mathbf{y}}_{\text{MM}}(\mathbf{f}_0) \\ \tilde{\mathbf{y}}_{\text{MM}}(\mathbf{f}_1) \\ \vdots \\ \tilde{\mathbf{y}}_{\text{MM}}(\mathbf{f}_i) \\ \vdots \\ \tilde{\mathbf{y}}_{\text{MM}}(\mathbf{f}_{N-1}) \end{bmatrix} \quad (\text{C.55})$$

with the $\text{KM} \times 1$ vector

$$\tilde{\mathbf{y}}_{\text{MM}}(\mathbf{f}_i) = \begin{bmatrix} \tilde{\mathbf{y}}(\mathbf{f}_i, \mathbf{f}'_0) \\ \tilde{\mathbf{y}}(\mathbf{f}_i, \mathbf{f}'_1) \\ \vdots \\ \tilde{\mathbf{y}}(\mathbf{f}_i, \mathbf{f}'_j) \\ \vdots \\ \tilde{\mathbf{y}}(\mathbf{f}_i, \mathbf{f}'_{M-1}) \end{bmatrix} \quad \text{for } i = 0, 1, \dots, N - 1 \quad (\text{C.56})$$

and the $\text{K} \times 1$ vector

$$\tilde{\mathbf{y}}(\mathbf{f}_i, \mathbf{f}'_j) = \begin{bmatrix} \tilde{\mathbf{y}}_0(\mathbf{f}_i, \mathbf{f}'_j) \\ \tilde{\mathbf{y}}_1(\mathbf{f}_i, \mathbf{f}'_j) \\ \vdots \\ \tilde{\mathbf{y}}_{\text{K}-1}(\mathbf{f}_i, \mathbf{f}'_j) \end{bmatrix} \quad \text{for } j = 0, 1, \dots, M - 1 \quad (\text{C.57})$$