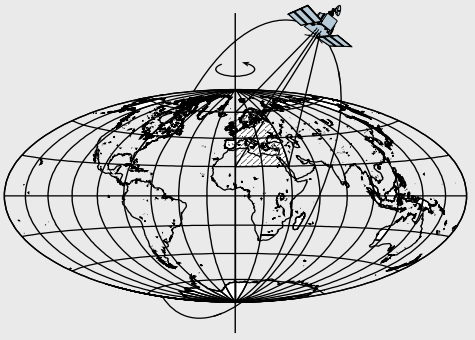
The Statistical Performance of the Matched Filter for Anomaly Detection Using Gravity Gradients

by

Christopher Jekeli and Tin Lian Abt



Report No. 494

Geodetic Science

The Ohio State University
Columbus, Ohio 43210

May 2010

The Statistical Performance of the Matched Filter for Anomaly Detection Using Gravity Gradients

by

Christopher Jekeli and Tin Lian Abt

Report 494

Division of Geodetic Science
School of Earth Sciences
Ohio State University

May 2010

Table of Contents

Abstract
This document first reviews the theory of detecting a subsurface linear anomaly using the matched filter applied to observations of the gravitational gradient in the presence of a nominal gravitational background field and along tracks crossing the anomaly orthogonally or at an arbitrary angle. The maximum filter output indicates the likely location along the track and, with appropriate statistical assumptions on the background field and measurement noise, it also serves as a test statistic in the probabilistic evaluation of the filter's performance. Different setups of the Neyman-Pearson statistical hypothesis test yield calculated probabilities of either a miss or a false alarm, respectively. The needed statistics of the maximum filter output are properly obtained using the distribution of order statistics. Through Monte Carlo simulations, we analyzed the ability of the matched filter to identify certain signals in typically correlated gravity fields using observations of elements of the gravity gradient tensor. We also evaluated the reliability of the hypothesis testing and of the associated calculated probabilities of misses and false alarms. We found that the hypothesis test that yields the probability of a miss is more robust than the one for the probability of a false alarm. Moreover, the probability of a miss is somewhat less than the probability of a false alarm under otherwise equal circumstances. Our simulations and statistical analyses confirm that the power of the tests increases as the signal strength increases and as more gradient tensor components per observation point are included. Finally, we found that the statistical methods apply only to single tracks (one-dimensional matched filter) and that the matched filter itself performs poorly for multiple parallel tracks (two-dimensional formulation) crossing a linear anomaly obliquely. Therefore, these methods are most useful, with respect to both the detection and the probability calculation, for single tracks of data crossing a linear anomaly (at arbitrary angle).

I Introduction

The problem of detecting a signal touches many disciplines, including communications, medicine, military operations, and image processing, among others. In simplest terms, the problem entails the identification of a known signal contained in background noise that itself has a known stochastic structure. In geophysics, the detection problem may be subsumed under the more general inverse problem of characterizing subsurface formations from measurements on or above the Earth's surface. Stable solutions to inverse problems require some form of constraint on the prediction errors, or on the parameters to be solved, or both. The matched filter (MF) is a particular type of solution where the essential constraint comes from the knowledge of the signal structure and the maximization of the signal-to-noise ratio.

The matched filter is not widely employed in geophysics (one example is the detection of sea mounts from satellite altimetry, White et al., 1983). The example studied in this report is the detection of a signal in the local ambient gravitational field using surface measurements of components of the gravity gradient tensor. The signal to be detected is due to the density contrast associated with a linear feature of either natural or anthropogenic origin, such as paleochannels, caves, large underground pipes, coal mines, tunnels, etc. The Federal Highway Administration (Meglich et al., 2005) recently sponsored an extensive study on the various measurement technologies available to detect lava tubes that could create a hazard or difficulty in highway construction and maintenance. Also, the USGS (Munk and Sheets, 1997) conducted a study of voids due to abandoned underground coal mines in Ohio that have led to the collapse of a major highway. Today's high-precision gravimeters are able to detect some of these anomalies, but gravity gradiometers are more sensitive to the very immediate mass anomalies and combining gradients in different directions potentially increases the detection capability (Butler, 1984; Romaides et al., 2001). Moreover, while other geophysical techniques such as magnetic, electromagnetic, and resistivity surveys, as well as ground penetrating radar are easier (cheaper) to employ for such detection, gravimetry is perhaps the most robust technique. Any of the methods based on electromagnetic fields or energies are highly susceptible to electromagnetic interferences from above the ground, which often are difficult to model, and to sub-surface dielectric and conductivity properties, which can vary by orders of magnitude. Gravimetry depends only on mass density, and we seek very specific anomalies in this distribution.

A well known feature of the matched filter is its connection to probabilistic measures of false alarms and missed detection. The maximum output of the filter, which indicates a candidate location of the signal being sought, is a test statistic of a standard Neyman-Pearson hypothesis test that under known probability distributions yields the probabilities of false alarms or of misses, depending on the setup of the test. Many of the details of the matched filter applied to gravity gradients, including comprehensive simulations for airborne measurements over extensive areas, have been elaborated by Dumrongchai (2007). Our aim is to review the essential mathematics and statistics of the matched filter and investigate in more detail the determination of these probabilities using Monte Carlo simulations to validate them.

The matched filter is developed first for a single track of one type of observation of a gravity gradient tensor element, then generalized to observations of multiple elements and on multiple parallel tracks. The subsurface anomaly is modeled as a right rectangular prism of constant negative density contrast, with its long, predominant dimension parallel to the plane containing the observations. First, we assume that it is orthogonal to the tracks, and then consider the case of an arbitrary azimuthal orientation. Completely arbitrary orientations in three angles can also

be formulated, although any rotation about the longitudinal axis of the prism will have little effect on the surface gradients, but a dip in the anomaly would clearly cause the signal strength at the surface to vary along the direction of this axis.

II Observations, Noise, and Signal

Let *z* be the observation at a point; let *n* be the noise; let *x* be the linear 1-D coordinate of the observation point; and let *s* be the sought signal due to a sub-surface anomaly. The observation model is then given by

$$z(x) = s(x - \bar{x}) + n(x), \tag{1}$$

where $\bar{x}$ is the unknown location of the signal relative to our imposed coordinate system. The number of evenly spaced observation points is *N*. The signal is of the same type as the observation (e.g., a gravitational gradient) and has a known structure; i.e., we can compute $s(x)$ based on known or estimated parameters of the anomaly. The noise, *n*, is assumed to be a stationary, correlated, random process with zero mean and a known covariance function, ϕ; however, it need not be a Gaussian process. Only when we attach probabilities to our detection process do we assume Gaussianity in the noise.

In practice the observation points are always discrete, and we suppose that the observations are distributed regularly along the (arbitrarily oriented) *x*-axis: $x_j = j\Delta x$, where Δx is a constant interval, and the range of index values is finite. We also write $z(x_j) = z_j$ and $n(x_j) = n_j$. The range of index values is chosen as follows (other possibilities could be used, e.g., $j = 0, \ldots, N-1$; see also the section on multiple tracks):

$$-\frac{N}{2} \le j \le \frac{N}{2} - 1, \tag{2}$$

where *N* is assumed to be a positive even integer. We are only interested in finding the location, $\bar{x}$, and we assume all other parameters of the anomaly (such as its depth) are given. Thus, letting $\bar{x} = x_m$, equation (1) becomes

$$z(x_j) = s(x_j - x_m) + n(x_j) \quad \text{or} \quad z_j = s_{j-m} + n_j. \tag{3}$$

In practice, the observations, z_j, are given together with the covariance function, ϕ, and the signal, $s(x)$. The only unknown is the location, x_m.

The noise comprises the gravitational background field, as well as instrument noise. The former may be approximated as a correlated, stationary field, while the latter typically is a white noise process (also stationary). The covariance function,

$$\phi = \phi_g + \phi_n, \tag{4}$$

is determined from an analysis of the local gravitational field, which yields a model for ϕ_g, and the known statistical information on the instrument noise, which gives ϕ_n.

In a Cartesian coordinate system, the gravitational gradients of a right rectangular prism of constant density and with its edges parallel to coordinate axes have straightforward analytic expressions. Appendix A derives these; and Appendix B shows how to compute the gradients for an arbitrarily oriented prism.

III The Matched Filter (Single Gradient, Single Track)

It is desired to find a filter that in its output identifies (a candidate for) the location of the signal if it is present. Like any filter, the matched filter is a convolution of the input (our observations), formulated for a *discrete* input as:

$$y(x_r) = \sum_{j=-N/2}^{N/2-1} h(x_r - x_j)\, z(x_j), \quad -\frac{N}{2} \le r \le \frac{N}{2} - 1, \tag{5}$$

where h is the filter function, which needs to be defined for x_k, $k = -N+1, \ldots, N-1$. Alternatively, with an appropriate definition of the observations outside its given limited domain, one can define the filter as follows (for infinite domains, $N \to \infty$, the two formulations would be identical):

$$y(x_r) = \sum_{j=-N/2}^{N/2-1} h(x_j)\, z(x_r - x_j), \quad -\frac{N}{2} \le r \le \frac{N}{2} - 1, \tag{6}$$

This form of the filter facilitates the derivation of its filter function, h. However, now the observations must be defined for $j = -N+1, \ldots, N-1$. This is accomplished by extending the observations either with zeros:

$$\tilde{z}_j = \begin{cases} 0, & -N+1 \le j < -N/2 \\ z_j, & -N/2 \le j \le N/2 - 1 \\ 0, & N/2 \le j \le N-1 \end{cases} \tag{7}$$

or, by duplicating the observations assuming they are periodic with period equal to the length of the observation track:

$$\tilde{z}_j = \begin{cases} z_{j+N}, & -N+1 \le j < -N/2 \\ z_j, & -N/2 \le j \le N/2 - 1 \\ z_{j-N}, & N/2 \le j \le N-1 \end{cases} \tag{8}$$

The latter option is a requirement when formulating the matched filter in terms of discrete Fourier transforms, and it is assumed here. However, we continue the derivations and all analyses in the space domain.

If the observations contain the signal, it would be identified best in the filter output where the signal-to-noise ratio is a maximum. Substituting the observations, modeled according to equation (3), into the filter (6), the output can be separated into a signal part and a noise part (using a simplified notation):

$$y_r = \sum_{j=-N/2}^{N/2-1} h_j s_{r-j-m} + \sum_{j=-N/2}^{N/2-1} h_j(x)\, n_{r-j}, \quad -\frac{N}{2} \le r \le \frac{N}{2}-1. \tag{9}$$

Consequently, we *define* the signal-to-noise ratio in terms of squared norms of these parts:

$$SNR_r = \frac{\left(\sum_{j=-N/2}^{N/2-1} h_j s_{r-j-m} \right)^2}{\mathcal{E}\left(\sum_{j=-N/2}^{N/2-1} h_j(x)\, n_{r-j} \right)^2}, \quad -\frac{N}{2} \le r \le \frac{N}{2}-1, \tag{10}$$

where the statistical expectation, $\mathcal{E}$, is used since we do not know the particular values of the noise.

Because the noise has zero mean and is stationary, its covariance function depends only on the difference between point coordinates. We *define* the order of differencing as the coordinate of the first quantity minus the coordinate of the second quantity in the covariance expression:

$$\begin{aligned} \operatorname{cov}\left(n(x_r - x_j), n(x_r - x_k)\right) &= \mathcal{E}\left(n(x_r - x_j)\, n(x_r - x_k)\right) \\ &= \phi\left((x_r - x_j) - (x_r - x_k)\right) \\ &= \phi\left(-x_j - (-x_k)\right) \end{aligned} \tag{11}$$

We denote by ϕ_- the $N \times N$ matrix of noise covariances that depend on the difference between two points, $-x_j$ and $-x_k$, and let $(\phi_-)_{j,k}$ be the individual elements. It is noted that the covariance matrix is symmetric (irrespective of the choice of order of differencing or the sign of the coordinates – but a different covariance matrix may result with a different choice of sign conventions– we just need to be consistent). Thus the denominator of the *SNR* becomes

$$\operatorname{denom}(SNR_r) = \sum_{j=-N/2}^{N/2-1} \sum_{k=-N/2}^{N/2-1} h_j (\phi_-)_{j,k} h_k, \quad \forall r, \tag{12}$$

and the denominator, therefore, does not dependent on *r*.

For a more compact notation, let the N filter-function values for the individual points be collected in an $N \times 1$ vector:

$$\boldsymbol{h}^{\mathrm{T}} = \left(h_{-N/2}, h_{-N/2+1}, \ldots, h_{N/2-1} \right). \tag{13}$$

Clearly, the denominator of *SNR* can now be written as

$$\mathrm{denom}\left(SNR_r \right) = \boldsymbol{h}^T \phi_{\_} \boldsymbol{h}, \quad \forall r \,, \tag{14}$$

and the numerator is

$$\mathrm{num}\left(SNR_r \right) = \left(\boldsymbol{h}^T \boldsymbol{s}_{r,m} \right)^2, \quad \forall r \,, \tag{15}$$

where

$$\boldsymbol{s}_{r,m}^{\mathrm{T}} = \left(s_{r+N/2-m}, \ldots, s_{r-(N/2-1)-m} \right). \tag{16}$$

The signal must be defined outside the observation domain; and, we use zeros since the anomaly, in any case, is assumed to generate a highly localized signal.

We wish to find an expression for the filter function that maximizes the *SNR* at $x_r = x_m$. From equation (10), the *SNR* does not depend on the scale of the filter function, and since the denominator is independent of x_r, we may insist that for whatever filter function, the denominator is *constant*, say c^2. Thus, the maximization of the *SNR* can be recast as a constrained maximization problem: maximize the numerator subject to the condition that the denominator is c^2. We set up a cost function with Lagrange multiplier:

$$\chi = \left(\boldsymbol{h}^{\mathrm{T}} \boldsymbol{s}_{\_} \right)^2 + \lambda^2 \left(c^2 - \boldsymbol{h}^{\mathrm{T}} \phi_{\_} \boldsymbol{h} \right), \tag{17}$$

where we have dispensed with the subscript notation on the signal since we consider maximizing the *SNR* at the location of the signal; that is, the signal in equation (17) is the signal in equation (16) with $r = m$:

$$\boldsymbol{s}_{m,m}^{\mathrm{T}} = \left(s_{N/2}, \ldots, s_{-N/2+1} \right) = \boldsymbol{s}_{\_}^{\mathrm{T}}, \tag{18}$$

again, denoting the reversal of signs in the coordinates with the subscripted minus sign. Taking derivatives of χ with respect to the unknowns, $\boldsymbol{h}$ and λ^2, we find:

$$\frac{\partial \chi}{\partial \boldsymbol{h}} = 2\boldsymbol{h}^{\mathrm{T}} \boldsymbol{s}_{\_} \boldsymbol{s}_{\_}^{\mathrm{T}} - \lambda^2 \left(\boldsymbol{h}^{\mathrm{T}} \phi_{\_} + \boldsymbol{h}^{\mathrm{T}} \phi_{\_}^{\mathrm{T}} \right), \tag{19}$$

$$\frac{\partial \chi}{\partial\left(\lambda^2\right)} = c^2 - \boldsymbol{h}^{\mathrm{T}} \phi_{\_} \boldsymbol{h}. \tag{20}$$

Since the covariance matrix is symmetric, we obtain, upon setting the derivatives to zero:

$$\boldsymbol{h}^{\mathrm{T}} \boldsymbol{s}_{\_} \boldsymbol{s}_{\_}^{\mathrm{T}} = \lambda^2 \boldsymbol{h}^{\mathrm{T}} \phi_{\_}, \tag{21}$$

$$\boldsymbol{h}^{\mathrm{T}} \phi_{\_} \boldsymbol{h} = c^2. \tag{22}$$

These are necessary conditions for the maximum of χ subject to the constraint. It is a standard procedure (not given here) to show that they are also sufficient conditions. Note that equation (22) is simply a repeat of our constraint. Post-multiplying equation (21) by $\boldsymbol{h}$, we find:

$$\lambda^2 = \frac{\boldsymbol{h}^{\mathrm{T}} \boldsymbol{s}_{\_} \boldsymbol{s}_{\_}^{\mathrm{T}} \boldsymbol{h}}{c^2}, \tag{23}$$

which is, in fact, the maximum *SNR*. Taking the transpose of equation (21) and noting that the covariance matrix is invertible, we have

$$\phi_{\_}^{-1} \boldsymbol{s}_{\_} \boldsymbol{s}_{\_}^{\mathrm{T}} \boldsymbol{h} = \lambda^2 \boldsymbol{h}. \tag{24}$$

The matrix, $\phi_{\_}^{-1} \boldsymbol{s}_{\_} \boldsymbol{s}_{\_}^{\mathrm{T}}$, has rank 1; and, therefore, all but one of its eigenvalues are zero. From the above equation, which is in the form that identifies all the eigenvectors, $\boldsymbol{h}$, we seek the one corresponding to the non-zero eigenvalue (i.e., we assume $\lambda^2 \neq 0$). It is easily verified that

$$\boldsymbol{h} = \frac{c}{\sqrt{\boldsymbol{s}_{\_}^{\mathrm{T}} \phi_{\_}^{-1} \boldsymbol{s}_{\_}}} \phi_{\_}^{-1} \boldsymbol{s}_{\_} \tag{25}$$

and

$$\begin{aligned} \lambda^2 &= \boldsymbol{s}_{\_}^{\mathrm{T}} \phi_{\_}^{-1} \boldsymbol{s}_{\_} \\ &= \sum_{j=-N/2}^{N/2-1} \sum_{k=-N/2}^{N/2-1} s_{-j} \left(\phi_{\_}^{-1}\right)_{j,k} s_{-k} \end{aligned} \tag{26}$$

satisfy equations (23) and (24).

Recalling that the scale of the filter function does not influence the detection capability of the matched filter, we may define $c^2 = 1$, which implies that

$$\boldsymbol{h} = (1/\lambda) \phi_{\_}^{-1} \boldsymbol{s}_{\_} \quad \Rightarrow \quad h_j = \frac{1}{\lambda} \sum_{k=-N/2}^{N/2-1} \left(\phi_{\_}^{-1}\right)_{j,k} s_{-k}, \quad -\frac{N}{2} \leq j \leq \frac{N}{2} - 1, \tag{27}$$

which also makes the filter output unit-less. Note that in designing the filter we would never *assume* that the signal is absent (otherwise, there would be no reason to continue). Thus, the *SNR* as defined here is the one that would occur if the signal were present. If the signal, in fact, is not present, then we still calculate the *SNR* under the assumption that it is, but our matched filter will not be successful in locating the signal. To indicate that the filter function depends on a putative signal (which may or may not be the correct signal, and which, indeed, may not exist at all in the observations), we denote the modeled signal as $\overline{s}$, and the filter function and the maximum *SNR* become

$$\boldsymbol{h} = (1/\lambda)\phi_{-}^{-1}\overline{\boldsymbol{s}}_{-} \quad \text{or} \quad h_j = \frac{1}{\lambda}\sum_{k=-N/2}^{N/2-1}\left(\phi_{-}^{-1}\right)_{j,k}\overline{s}_{-k}, \quad -\frac{N}{2} \le j \le \frac{N}{2}-1; \tag{28}$$

$$\lambda^2 = \sum_{j=-N/2}^{N/2-1}\sum_{k=-N/2}^{N/2-1}\overline{s}_{-j}\left(\phi_{-}^{-1}\right)_{j,k}\overline{s}_{-k}\,. \tag{29}$$

Substituting equation (28) into equation (6) (with observations extended as in equations (7) or (8)), the filter output can be calculated as

$$y_r = \frac{1}{\lambda}\sum_{j=-N/2}^{N/2-1}\sum_{k=-N/2}^{N/2-1}\overline{s}_{-k}\left(\phi_{-}^{-1}\right)_{j,k}\tilde{z}_{r-j}, \quad -\frac{N}{2} \le r \le \frac{N}{2}-1\,. \tag{30}$$

For $r = -N/2$, the indices of $\tilde{z}$ are $r - j = 0, -1, \ldots, -N+1$. For $r = N/2-1$, the indices of $\tilde{z}$ are $r - j = N-1, N-2, \ldots, 0$. Therefore, the observations need to be extended as in equations (7) or (8). The (given) signal, $\overline{s}_j$, is defined as if it were located at $x = 0$ (or symmetric with respect to $x = 0$), with $j = -N/2, \ldots, N/2$ (although the first point is not used). Finally, it is noted that the filter output is a scalar quantity, which can be written in vector-matrix notation as

$$y_r = \boldsymbol{h}^{\mathrm{T}}\tilde{\boldsymbol{z}}_{-}^{(r)}, \quad -\frac{N}{2} \le r \le \frac{N}{2}-1, \tag{31}$$

where $\boldsymbol{h} = \dfrac{1}{\lambda}\phi_{-}^{-1}\overline{\boldsymbol{s}}_{-}$, and where

$$\overline{\boldsymbol{s}}_{-}^{\mathrm{T}} = \left(\overline{s}_{N/2}, \overline{s}_{N/2-1}, \ldots, \overline{s}_{-N/2+1}\right)^{\mathrm{T}} \tag{32}$$

and

$$\tilde{\boldsymbol{z}}_{-}^{(r)} = \left(\tilde{z}_{N/2+r}, \tilde{z}_{N/2-1+r}, \ldots, \tilde{z}_{-N/2+1+r}\right)^{\mathrm{T}}. \tag{33}$$

The maximum output of the filter indicates a candidate location of the signal; and we denote the maximum as

$$\bar{y} = \max_r \left(y_r \right). \tag{34}$$

IV Multiple Observations

The matched filter developed in the previous section is easily extended to multiple observations per point along the track. Instead of scalar functions, $z(x)$, $s(x)$, we now have vector functions, $\boldsymbol{z}(x) = \left(z1(x), z2(x), \ldots \right)^{\mathrm{T}}$, $\boldsymbol{s}(x) = \left(s1(x), s2(x), \ldots \right)^{\mathrm{T}}$, containing the particular gradients (or combinations thereof) at each point. Note that the observation points are still along a single, straight track. The matched filter function is given as in equation (28) by

$$\boldsymbol{h} = \left(1/\lambda \right) \phi_{-}^{-1} \bar{\boldsymbol{s}}_{-}, \tag{35}$$

where the signal-to-noise ratio is still given by equation (26), but now

$$\bar{\boldsymbol{s}}_{-} = \left(\bar{s}1_{N/2}, \bar{s}2_{N/2}, \ldots, \bar{s}1_{-N/2+1}, \bar{s}2_{-N/2+1}, \ldots \right)^{\mathrm{T}}, \tag{36}$$

and

$$\phi_{-} = \left(\begin{array}{ccc|c|ccc} \mathrm{cov}\left(n1_{N/2}, n1_{N/2}\right) & \mathrm{cov}\left(n1_{N/2}, n2_{N/2}\right) & \cdots & & \mathrm{cov}\left(n1_{N/2}, n1_{-N/2+1}\right) & \mathrm{cov}\left(n1_{N/2}, n2_{-N/2+1}\right) & \cdots \\ \mathrm{cov}\left(n2_{N/2}, n1_{N/2}\right) & \mathrm{cov}\left(n2_{N/2}, n2_{N/2}\right) & \cdots & \cdots & \mathrm{cov}\left(n2_{N/2}, n1_{-N/2+1}\right) & \mathrm{cov}\left(n2_{N/2}, n2_{-N/2+1}\right) & \cdots \\ \vdots & \vdots & \ddots & & \vdots & \vdots & \ddots \\ \hline \vdots & & \ddots & & \vdots & & \\ \hline \mathrm{cov}\left(n1_{-N/2+1}, n1_{N/2}\right) & \mathrm{cov}\left(n1_{-N/2+1}, n2_{N/2}\right) & \cdots & & \mathrm{cov}\left(n1_{-N/2+1}, n1_{-N/2+1}\right) & \mathrm{cov}\left(n1_{-N/2+1}, n2_{-N/2+1}\right) & \cdots \\ \mathrm{cov}\left(n2_{-N/2+1}, n1_{N/2}\right) & \mathrm{cov}\left(n2_{-N/2+1}, n2_{N/2}\right) & \cdots & \cdots & \mathrm{cov}\left(n2_{-N/2+1}, n1_{-N/2+1}\right) & \mathrm{cov}\left(n2_{-N/2+1}, n2_{-N/2+1}\right) & \cdots \\ \vdots & \vdots & \ddots & & \vdots & \vdots & \ddots \end{array} \right) \tag{37}$$

The MF output, given by equation (31), again, is a scalar, and the observation vector (equation (33)) in this case is

$$\tilde{\boldsymbol{z}}_{-}^{(r)} = \left(\tilde{z}1_{N/2+r}, \tilde{z}2_{N/2+r}, \ldots, \tilde{z}1_{N/2-1+r}, \tilde{z}2_{N/2-1+r}, \ldots, \tilde{z}1_{-N/2+1+r}, \tilde{z}2_{-N/2+1+r}, \ldots \right)^{\mathrm{T}}. \tag{38}$$

In case observations are linear combinations of gradients, for example,

$$\Delta\Gamma = \Gamma_{22} - \Gamma_{11} = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} \Gamma_{22} \\ \Gamma_{11} \end{pmatrix}, \tag{39}$$

then it is only necessary to ensure that the correct covariance element is computed for this combination, and between this combination and possibly other gradients. For the example above, with a second observation of $2\Gamma_{12}$, we have

$$\begin{aligned}\operatorname{cov}(\Delta\Gamma,\Delta\Gamma) &= \begin{pmatrix}1 & -1\end{pmatrix}\begin{pmatrix}\operatorname{cov}(\Gamma_{22},\Gamma_{22}) & \operatorname{cov}(\Gamma_{22},\Gamma_{11})\\ \operatorname{cov}(\Gamma_{11},\Gamma_{22}) & \operatorname{cov}(\Gamma_{11},\Gamma_{11})\end{pmatrix}\begin{pmatrix}1\\-1\end{pmatrix} \\ &= \operatorname{cov}(\Gamma_{22},\Gamma_{22}) - \operatorname{cov}(\Gamma_{11},\Gamma_{22}) - \operatorname{cov}(\Gamma_{22},\Gamma_{11}) + \operatorname{cov}(\Gamma_{11},\Gamma_{11})\end{aligned} \tag{40}$$

$$\operatorname{cov}(\Delta\Gamma,2\Gamma_{12}) = \operatorname{cov}(\Gamma_{22},2\Gamma_{12}) - \operatorname{cov}(\Gamma_{11},2\Gamma_{12}). \tag{41}$$

Then, the covariance matrix, equation (37), becomes

$$\phi_{-} = \left(\begin{array}{cc|c|cc}
\operatorname{cov}\left(2\Gamma_{12_{N/2}},2\Gamma_{12_{N/2}}\right) & \operatorname{cov}\left(2\Gamma_{12_{N/2}},\Delta\Gamma_{N/2}\right) & \cdots & \operatorname{cov}\left(2\Gamma_{12_{N/2}},2\Gamma_{12_{-N/2+1}}\right) & \operatorname{cov}\left(2\Gamma_{12_{N/2}},\Delta\Gamma_{-N/2+1}\right)\\
\operatorname{cov}\left(\Delta\Gamma_{N/2},2\Gamma_{12_{N/2}}\right) & \operatorname{cov}\left(\Delta\Gamma_{N/2},\Delta\Gamma_{N/2}\right) & \cdots & \operatorname{cov}\left(\Delta\Gamma_{N/2},2\Gamma_{12_{-N/2+1}}\right) & \operatorname{cov}\left(\Delta\Gamma_{N/2},\Delta\Gamma_{-N/2+1}\right)\\
\hline
\vdots & \vdots & \ddots & \vdots & \vdots\\
\hline
\operatorname{cov}\left(2\Gamma_{12_{-N/2+1}},2\Gamma_{12_{N/2}}\right) & \operatorname{cov}\left(2\Gamma_{12_{-N/2+1}},\Delta\Gamma_{N/2}\right) & \cdots & \operatorname{cov}\left(2\Gamma_{12_{-N/2+1}},2\Gamma_{12_{-N/2+1}}\right) & \operatorname{cov}\left(2\Gamma_{12_{-N/2+1}},\Delta\Gamma_{-N/2+1}\right)\\
\operatorname{cov}\left(\Delta\Gamma_{-N/2+1},2\Gamma_{12_{N/2}}\right) & \operatorname{cov}\left(\Delta\Gamma_{-N/2+1},\Delta\Gamma_{N/2}\right) & \cdots & \operatorname{cov}\left(\Delta\Gamma_{-N/2+1},2\Gamma_{12_{-N/2+1}}\right) & \operatorname{cov}\left(\Delta\Gamma_{-N/2+1},\Delta\Gamma_{-N/2+1}\right)
\end{array}\right) \tag{42}$$

V Multiple Tracks

In Section III, we derived the formula of the matched filter for observations along a single track. This can be generalized to multiple parallel tracks (or a two-dimensional grid of observations) by realizing that the filter is then a two-dimensional convolution of two-dimensional observations and a two-dimensional filter function: To provide a slightly different indexing scheme, let us assume that the number of tracks and the number of observation points per track are both odd integers (the formulations for even integers would be very similar, as illustrated for the single-track case). From equations (28) and (30), we have for one track (and odd *N*)

$$y_r = \sum_{j=-(N-1)/2}^{(N-1)/2} h_j \tilde{z}_{r-j}, \quad -\frac{N-1}{2} \le r \le \frac{N-1}{2}. \tag{43}$$

Extending this to a 2-D convolution then yields

$$y_{r_1,r_2} = \sum_{j_1=-(N_1-1)/2}^{(N_1-1)/2} \sum_{j_2=-(N_2-1)/2}^{(N_2-1)/2} h_{j_1,j_2} \tilde{z}_{r_1-j_1,r_2-j_2}, \quad -\frac{N_1-1}{2} \le r_1 \le \frac{N_1-1}{2}, \quad -\frac{N_2-1}{2} \le r_2 \le \frac{N_2-1}{2}. \tag{44}$$

Let the first subscript denote the along-track coordinate and the second subscript the cross-track coordinate. The filter function is now (analogous to equation (28))

$$h_{j_1,j_2}=\frac{1}{\lambda}\sum_{k_1=-(N_1-1)/2}^{(N_1-1)/2}\sum_{k_2=-(N_2-1)/2}^{(N_2-1)/2}\phi_{-}^{-1}\left(j_1,k_1,j_2,k_2\right)\overline{s}_{-k_1,-k_2},\quad -\frac{N_1-1}{2}\le j_1\le\frac{N_1-1}{2},\quad -\frac{N_2-1}{2}\le j_2\le\frac{N_2-1}{2} \tag{45}$$

where $\phi_{-}^{-1}\left(j_1,k_1,j_2,k_2\right)$ is the appropriate element of the inverse of the covariance matrix:

$$\phi_{-}=\left[\operatorname{cov}\left(n_{-j_1,-j_2},n_{-k_1,-k_2}\right)\right]. \tag{46}$$

It is noted that for each $\left(r_1,r_2\right)$ the double sum in equation (44) is a scalar, just like in the single-track case; that is, the filter output is still a scalar function.

In order to achieve a more compact notation using vectors, we may concatenate the tracks and duly account for the covariances of the background noises between points of different tracks. We write

$$\tilde{\boldsymbol{z}}_{-}^{(r_1,r_2)}=\begin{pmatrix}\tilde{z}_{(N_1-1)/2+r_1,(N_2-1)/2+r_2},\tilde{z}_{(N_1-1)/2-1+r_1,(N_2-1)/2+r_2},\ldots,\tilde{z}_{-(N_1-1)/2+r_1,(N_2-1)/2+r_2},\ldots,\\ \tilde{z}_{(N_1-1)/2+r_1,-(N_2-1)/2+r_2},\tilde{z}_{(N_1-1)/2-1+r_1,-(N_2-1)/2+r_2},\ldots,\tilde{z}_{-(N_1-1)/2+r_1,-(N_2-1)/2+r_2}\end{pmatrix}^{\mathrm{T}}, \tag{47}$$

where each color corresponds to a track. Also,

$$\overline{\boldsymbol{s}}_{-}^{\mathrm{T}}=\begin{pmatrix}\overline{s}_{(N_1-1)/2,(N_2-1)/2},\overline{s}_{(N_1-1)/2-1,(N_2-1)/2},\ldots,\overline{s}_{-(N_1-1)/2,(N_2-1)/2},\ldots,\\ \overline{s}_{(N_1-1)/2,-(N_2-1)/2},\overline{s}_{(N_1-1)/2-1,-(N_2-1)/2},\ldots,\overline{s}_{-(N_1-1)/2,-(N_2-1)/2}\end{pmatrix}^{\mathrm{T}}; \tag{48}$$

and

$$\phi_{-}=\left[\operatorname{cov}\left(n_{-j_1,-j_2},n_{-k_1,-k_2}\right)\right]=\begin{pmatrix}C_{\frac{N_2-1}{2},\frac{N_2-1}{2}} & \cdots & C_{\frac{N_2-1}{2},-\frac{N_2-1}{2}}\\ \vdots & \ddots & \vdots\\ C_{-\frac{N_2-1}{2},\frac{N_2-1}{2}} & \cdots & C_{-\frac{N_2-1}{2},-\frac{N_2-1}{2}}\end{pmatrix}_{N_1N_2\times N_1N_2}, \tag{49}$$

where the dimensions of the matrix are indicated in the subscript, and where

$$C_{\frac{N_2-1}{2},\frac{N_2-1}{2}}=\begin{pmatrix}\left(\frac{N_1-1}{2}-\frac{N_1-1}{2},\frac{N_2-1}{2}-\frac{N_2-1}{2}\right) & \cdots & \left(\frac{N_1-1}{2}+\frac{N_1-1}{2},\frac{N_2-1}{2}-\frac{N_2-1}{2}\right)\\ \vdots & \ddots & \vdots\\ \left(-\frac{N_1-1}{2}-\frac{N_1-1}{2},\frac{N_2-1}{2}-\frac{N_2-1}{2}\right) & \cdots & \left(-\frac{N_1-1}{2}+\frac{N_1-1}{2},\frac{N_2-1}{2}-\frac{N_2-1}{2}\right)\end{pmatrix}_{N_1\times N_1} \tag{50}$$

denotes the sub-matrix of covariances among noises on the $k=-\left(N_2-1\right)/2$ track,

$$C_{-\frac{N_2-1}{2},\frac{N_2-1}{2}}=\begin{pmatrix}\left(\frac{N_1-1}{2}-\frac{N_1-1}{2},-\frac{N_2-1}{2}-\frac{N_2-1}{2}\right) & \cdots & \left(\frac{N_1-1}{2}+\frac{N_1-1}{2},-\frac{N_2-1}{2}-\frac{N_2-1}{2}\right)\\ \vdots & \ddots & \vdots\\ \left(-\frac{N_1-1}{2}-\frac{N_1-1}{2},-\frac{N_2-1}{2}-\frac{N_2-1}{2}\right) & \cdots & \left(-\frac{N_1-1}{2}+\frac{N_1-1}{2},-\frac{N_2-1}{2}-\frac{N_2-1}{2}\right)\end{pmatrix}_{N_1\times N_1} \tag{51}$$

denotes the sub-matrix of covariances between the $k=(N_2-1)/2$ track and the $k=-(N_2-1)/2$ track,

$$C_{\frac{N_2-1}{2},-\frac{N_2-1}{2}}=\begin{pmatrix}\left(\frac{N_1-1}{2}-\frac{N_1-1}{2},\frac{N_2-1}{2}+\frac{N_2-1}{2}\right) & \cdots & \left(\frac{N_1-1}{2}+\frac{N_1-1}{2},\frac{N_2-1}{2}+\frac{N_2-1}{2}\right)\\ \vdots & \ddots & \vdots\\ \left(-\frac{N_1-1}{2}-\frac{N_1-1}{2},\frac{N_2-1}{2}+\frac{N_2-1}{2}\right) & \cdots & \left(-\frac{N_1-1}{2}+\frac{N_1-1}{2},\frac{N_2-1}{2}+\frac{N_2-1}{2}\right)\end{pmatrix}_{N_1\times N_1} \tag{52}$$

denotes the sub-matrix of covariances between the $k=-(N_2-1)/2$ track and the $k=(N_2-1)/2$ track, and

$$C_{-\frac{N_2-1}{2},-\frac{N_2-1}{2}}=\begin{pmatrix}\left(\frac{N_1-1}{2}-\frac{N_1-1}{2},-\frac{N_2-1}{2}+\frac{N_2-1}{2}\right) & \cdots & \left(\frac{N_1-1}{2}+\frac{N_1-1}{2},-\frac{N_2-1}{2}+\frac{N_2-1}{2}\right)\\ \vdots & \ddots & \vdots\\ \left(-\frac{N_1-1}{2}-\frac{N_1-1}{2},-\frac{N_2-1}{2}+\frac{N_2-1}{2}\right) & \cdots & \left(-\frac{N_1-1}{2}+\frac{N_1-1}{2},-\frac{N_2-1}{2}+\frac{N_2-1}{2}\right)\end{pmatrix}_{N_1\times N_1} \tag{53}$$

denotes the sub-matrix of covariances among noises on the $k=(N_2-1)/2$ track. In each matrix, the element $\left(-j_1-(-j_2),-k_1-(-k_2)\right)$ denotes the differences in coordinates that are used to evaluate the covariance function (always the first coordinate minus the second coordinate). The element of ϕ_- is located by the subscripts $(N_1j_2+j_1,N_1k_2+k_1)$, allowing subscripts to range over the intervals as in equations (44) and (45).

With these definitions, we can write as before (see equations (26) and (31) with (28)):

$$\lambda^2=\overline{\mathbf{s}}_-^{\mathrm{T}}\phi_-^{-1}\overline{\mathbf{s}}_-, \tag{54}$$

and

$$y_{r_1,r_2}=\frac{1}{\lambda}\overline{\mathbf{s}}_-^{\mathrm{T}}\phi_-^{-1}\tilde{\mathbf{z}}_-^{(r_1,r_2)},\quad -\frac{N_1-1}{2}\le r_1\le\frac{N_1-1}{2},\quad -\frac{N_2-1}{2}\le r_2\le\frac{N_2-1}{2}. \tag{55}$$

For multiple observations at each point, it is necessary only to order these appropriately in the observation and signal vectors, and to substitute the appropriate covariance matrices.

It is noted that the absolute *maximum* of the filter output obviously refers to a single point in the observation grid and this would be suitable to identify a local (monopole-like) signal. However, in principle, the 2-D matched filter should be able to identify any 2-D signal at the location of the maximum output value on the observation grid.

By the symmetry in the indices with respect to the origin (since the numbers of the observation points and of the tracks are odd), no additional signal values need to be defined; however, the observations need to be extended for the following indices:

$$\tilde{z}_{j_1,j_2} = \begin{cases} z_{j_1+N_1,j_2}, & -N_1 \le j_1 < -(N_1-1)/2, \quad -(N_2-1)/2 \le j_2 \le (N_2-1)/2 \\ z_{j_1-N_1,j_2}, & (N_1-1)/2 < j_1 \le N_1, \quad -(N_2-1)/2 \le j_2 \le (N_2-1)/2 \\ z_{j_1,j_2}, & -(N_1-1)/2 \le j_1 \le (N_1-1)/2, \quad -(N_2-1)/2 \le j_2 \le (N_2-1)/2 \\ \tilde{z}_{j_1,j_2+N_2}, & -N_1 \le j_1 < N_1, \quad -N_2 \le j_2 < -(N_2-1)/2 \\ \tilde{z}_{j_1,j_2-N_2}, & -N_1 \le j_1 < N_1, \quad (N_2-1)/2 < j_2 \le N_2 \end{cases} \tag{56}$$

VI Statistical Performance

Once the maximum filter output has been determined from the given data, we may wish to assign a probability of false alarm or a probability of a miss. We propose two hypothesis setups in order to assign these probabilities. Under the first setup, Setup A, we have the following null hypothesis and its alternative:

Setup A

$H_0^{(A)}$: there is no (detectable) signal in the observations. (57)

$H_1^{(A)}$: there is a signal in the observations. (58)

We assume that these hypotheses represent two mutually exclusive and exhaustive events, where $H_0^{(A)}$ implies that either the signal does not exist or it is too deeply buried in the noise to be detectable – for our purposes these two scenarios are equivalent. The other setup, Setup B, defines the null hypothesis as supposing that the signal exists in the observations.

Setup B:

$H_0^{(B)}$: there is a signal in the observations. (59)

$H_1^{(B)}$: there is no (detectable) signal in the observations. (60)

We employ the Neyman-Pearson tests, which produce the probability of a Type II error given the probability of a Type I error. An error of Type I occurs when the data wrongly lead to a rejection of the null hypothesis; and, a Type II error occurs when the data wrongly accept the null hypothesis. Therefore, the Setup A leads to the determination of the probability of a miss, and the Setup B leads to the probability of a false alarm, in each case for a given acceptable level of Type I error.

It can be shown (Appendix C) that the maximum filter output, $\bar{y}$ (equation (34)), which indicates a candidate location of the signal, is also a *test statistic* for the hypotheses of Setups A and B. Therefore, in order to compute the required probabilities, we require the *probability density function* (pdf) of $\bar{y}$. Toward this end, we first determine the pdf of the MF output, in general. This can be computed if the pdf of the noise, *n*, is known; in fact, it is assumed to be Gaussian with zero mean and covariance function, ϕ.

We consider only the case of a single track (and later comment on the situation with multiple tracks). The pdf of the MF output, being a linear combination of the observations (and hence the noise), is also Gaussian. The mean (or expectation) is given by (assuming even *N*, but it makes no essential difference if *N* is odd)

$$\mathcal{E}\left(y_r\right)=\frac{1}{\lambda}\sum_{j=-N/2}^{N/2-1}\sum_{k=-N/2}^{N/2-1}\bar{s}_{-k}^{\mathrm{T}}\left(\phi_{-}^{-1}\right)_{j,k}s_{r-j-m},\quad -\frac{N}{2}\le r\le\frac{N}{2}-1, \tag{61}$$

where, again, $\bar{s}$ is the assumed signal and *s* is the actual signal (which may differ from the assumed signal due to errors in various parameters, such as dimensions, depth (including infinite depth), and orientation of the anomaly). The expected value of the *maximum* filter output thus depends on the hypothesis. If there is no signal ($s=0$), then the expectation is zero under either $H_0^{(A)}$ or $H_1^{(B)}$. If the assumed signal exists, $s=\bar{s}$ (either $H_0^{(B)}$ or $H_1^{(A)}$ is true), then the expectation of the *maximum* filter output is the square root of the signal-to-noise ratio, λ, which is obtained by substituting observations (3) into filter (30), setting $r=m$, and comparing to equation (26). We have:

$$\mathcal{E}\left(\bar{y}\right)=\begin{cases}0, & H_0^{(A)}\text{ is true, or }H_1^{(B)}\text{ is true}\\ \lambda, & H_0^{(B)}\text{ is true, or }H_1^{(A)}\text{ is true}\end{cases} \tag{62}$$

The variance of the filter output for all *r* (including its maximum) is unity, irrespective of a signal in the observations. This follows almost immediately from equation (31):

$$\mathrm{var}\left(y_r\right)=1,\quad \forall r\,; \tag{63}$$

see also the derivations leading to equation (D.9) in Appendix D.

Now, while the pdf of the MF output, y_r, is Gaussian, that of the *maximum* output, $\bar{y}$, is not necessarily Gaussian. If the signal in the observations is strong, the maximum filter output does not vary in location with different realization of the observations, and we may expect that $\bar{y}$ is Gaussian with unit variance and mean equal to λ. But, if the signal in the observations is virtually non-existent, the maximum output is located randomly and its pdf more closely assumes the density corresponding to an *order* statistic, that is, the density of a variable that has a specific hierarchical place among a sample of given size.

From the study of order statistics (e.g., David, 1981), the maximum, $\bar{y}$, has a well defined probability density function, $f_{\bar{y}}$, if the filter outputs, y_r, are identically and independently distributed for all *r*:

$$f_{\bar{y}}(\bar{y}) = N\left(F_y(\bar{y})\right)^{N-1} f_y(\bar{y}), \tag{64}$$

where N is the number of points (samples), f_y is the probability density function of y_r, and F_y is its cumulative distribution function. From equations (61) and (63), the MF outputs, y_r, in the absence of a signal are identically distributed (zero-mean, unit-variance Gaussian). If the background noise is uncorrelated and the signal is like an impulse function, then the MF output is easily seen to be also uncorrelated (Appendix D). However, in the more general case of either correlated noise and/or a more extended signal, the output is only approximately uncorrelated, depending on the amount of correlation of the noise and extension of the signal. Nevertheless, assuming that this approximation is adequate, we model the distribution of $\bar{y}$ under hypothesis $H_0^{(A)}$ (no signal) according to the density given by equation (64).

Thus, we assume that under hypothesis $H_1^{(A)}$ the probability density of $\bar{y}$ is Gaussian:

$$f_{\bar{y}|H_1^{(A)}}\left(\bar{y}\middle|H_1^{(A)}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\bar{y}-\lambda)^2}; \tag{65}$$

since its mean is λ (equation (62)), and the variance is unity (equation (63)). And, under hypothesis $H_0^{(A)}$, we assume that $\bar{y}$ is distributed according to the density given by equation (64) with

$$f_{y|H_0^{(A)}}\left(\bar{y}\middle|H_0^{(A)}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\bar{y}^2}, \tag{66}$$

and

$$F_{y|H_0^{(A)}}\left(\bar{y}\middle|H_0^{(A)}\right) = \frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{\bar{y}}{\sqrt{2}}\right). \tag{67}$$

where erf is the error function. That is, if there is no signal in the observations ($s = 0$), the mean of the filter output is zero and the variance is still unity. We write

$$f_{\bar{y}|H_0^{(A)}}\left(\bar{y}\middle|H_0^{(A)}\right) = \frac{N}{\sqrt{2\pi}}\left(\frac{1}{2} + \frac{1}{2}\operatorname{erf}\left(\frac{\bar{y}}{\sqrt{2}}\right)\right)^{N-1} e^{-\frac{1}{2}\bar{y}^2}. \tag{68}$$

Note that the mean and variance of the *maximum* filter output are not those of the general filter output, y_r. Under Setup B, these probability densities are reversed in accordance with the reversal of the hypotheses:

$$f_{\bar{y}|H_0^{(B)}}\left(\bar{y}\middle|H_0^{(B)}\right)=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(\bar{y}-\lambda)^2}, \tag{69}$$

$$f_{\bar{y}|H_1^{(B)}}\left(\bar{y}\middle|H_1^{(B)}\right)=\frac{N}{\sqrt{2\pi}}\left(\frac{1}{2}+\frac{1}{2}\operatorname{erf}\left(\frac{\bar{y}}{\sqrt{2}}\right)\right)^{N-1}e^{-\frac{1}{2}\bar{y}^2}. \tag{70}$$

Figure 1 exhibits the non-symmetric character of the density for the maximum output when no signal is present in the observations, compared to the usual symmetry of the Gaussian density for the case that the signal is detectable. This asymmetry influences the implementation as well as the performance of the statistical tests.

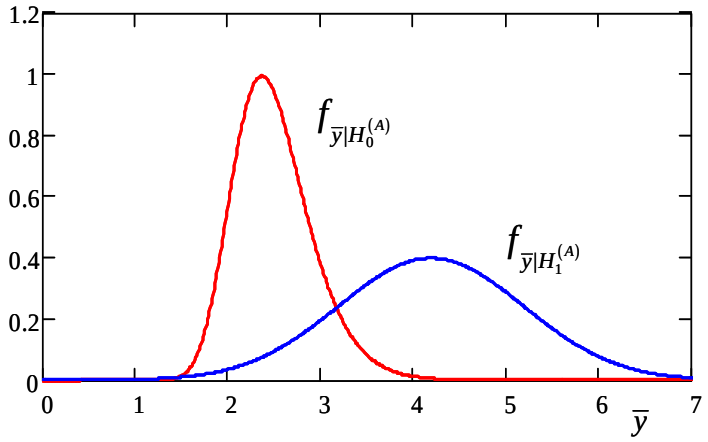


Figure 1: Probability densities for the maximum matched filter output under hypotheses of Setup A

The Neyman-Pearson hypothesis test is performed by first selecting a probability of Type I error, denoted α, that is acceptable. This is also the significance level of the test and it establishes a threshold, ψ_0, for the test statistic, $\bar{y}$. For Setups A and B, we have

$$\alpha^{(A)}=\int_{\psi_0^{(A)}}^{\infty}f_{\bar{y}|H_0^{(A)}}\left(\bar{y}\middle|H_0^{(A)}\right)d\bar{y}=\frac{N}{\sqrt{2\pi}}\int_{\psi_0^{(A)}}^{\infty}\left(\frac{1}{2}+\frac{1}{2}\operatorname{erf}\left(\frac{\bar{y}}{\sqrt{2}}\right)\right)^{N-1}e^{-\frac{1}{2}\bar{y}^2}d\bar{y}, \tag{71}$$

$$\alpha^{(B)}=\int_{-\infty}^{\psi_0^{(B)}}f_{\bar{y}|H_0^{(B)}}\left(\bar{y}\middle|H_0^{(B)}\right)d\bar{y}=F_{\bar{y}|H_0^{(B)}}\left(\psi_0^{(B)}\right)=\frac{1}{2}+\frac{1}{2}\operatorname{erf}\left(\frac{\psi_0^{(B)}-\lambda}{\sqrt{2}}\right), \tag{72}$$

The threshold, $\psi_0^{(A)}$, in Setup A must be determined numerically using equation (71) for the given $\alpha^{(A)}$; for Setup B, it is

$$\psi_0^{(B)} = \lambda + \sqrt{2}\,\mathrm{erf}^{-1}\left(2\alpha^{(B)} - 1\right). \tag{73}$$

If $\bar{y}$ is *greater than* the threshold, $\psi_0^{(A)}$, it falls in the critical region of the probability density function (one-tailed test) where it is considered to be statistically *inconsistent* with the null hypothesis; hence, we reject $H_0^{(A)}$ (that there is no signal in the observations) at the $\alpha^{(A)}$ significance level. If the output is less than the threshold, then we accept $H_0^{(A)}$. The inequalities are reversed for Setup B; $H_0^{(B)}$ is accepted if $\bar{y}$ is greater than the threshold, $\psi_0^{(B)}$.

The power of the test is defined as $p = 1 - \beta$, where β is the probability of making a Type II error (wrongly accepting the null hypothesis). For the two setups we have

$$\beta^{(A)} = \int_{-\infty}^{\psi_0^{(A)}} f_{\bar{y}|H_1^{(A)}}\left(\bar{y}\middle|H_1^{(A)}\right)d\bar{y} = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{\psi_0^{(A)} - \lambda}{\sqrt{2}}\right); \tag{74}$$

$$\beta^{(B)} = \int_{\psi_0^{(B)}}^{\infty} f_{\bar{y}|H_1^{(B)}}\left(\bar{y}\middle|H_1^{(B)}\right)d\bar{y} = \frac{N}{\sqrt{2\pi}}\int_{\psi_0^{(B)}}^{\infty}\left(\frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{\bar{y}}{\sqrt{2}}\right)\right)^{N-1} e^{-\frac{1}{2}\bar{y}^2}\,d\bar{y}\,. \tag{75}$$

The Type II error probability may be identified as the probability of a miss (false negative), $POM = \beta^{(A)}$, under Setup A, and as the probability of a false alarm (false positive), $POF = \beta^{(B)}$, under Setup B.

VII Numerical Examples of the MF

The matched filter as developed in the previous sections was tested using a simulated signal buried in background noise. As noted in Section I, the signal is due to an elongated right-rectangular prism orthogonal to the data track(s) and with sides parallel to the coordinate planes. Appendix A provides the appropriate formulas for the gravitational gradients. The "noise" consists of both a nominal gravitational field and instrument noise with given statistics. The gravitational field was generated from the power spectral density model of a typical gravity anomaly field extended to very high spatial frequency (1 cy/m) according to a power law, in keeping with the theory that the high-frequency gravitational field is generated primarily by the visible topography, which may be described as a fractal. In essence, the field is generated by a 2-D Fourier spectrum whose components are Gaussian random numbers that are scaled at each frequency by the square root of the psd at that frequency (Appendix E). The details of this type of psd modeling may be found in (Jekeli, 2003, 2010).

The results of two tests are shown in this section. The first test assumes observations of the vertical-vertical gradient, Γ_{33}, along the x_1-axis at a constant interval of $\Delta x_1 = 1$ m. The second test is based on observations of $2\Gamma_{12}$ and $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$ at these points. In both cases, the number of observation points is $N = 100$; and each observation includes instrument noise

(generated using a Gaussian random number generator) with standard deviation, $\sigma = 3$ E ($1\text{ E} = 10^{-9}\text{s}^{-2}$). The anomaly that generates the corresponding signals is a prism with density contrast, $\rho = -2670$ kg/m^3; along-track width, $a = 1$ m, cross-track length, $b = 100$ m; and height, $c = 2$ m, where the depth to the top of the prism is $d = 2$ m. The geometric center of the anomaly is located at $x_1 = -20$ m, $x_2 = 0$ m, $x_3 = -3$ m. Figures 2-4 show the profiles of the observations, z_j, and the signals due to the anomaly in each case. Note that in this example the signal gradients, Γ_{12}, Γ_{22}, are vanishing small (not exactly zero due to the finite length of the anomaly). Therefore, because of Laplace's constraint, we have for the signal, $\Gamma_{33} \approx \Delta\Gamma$.

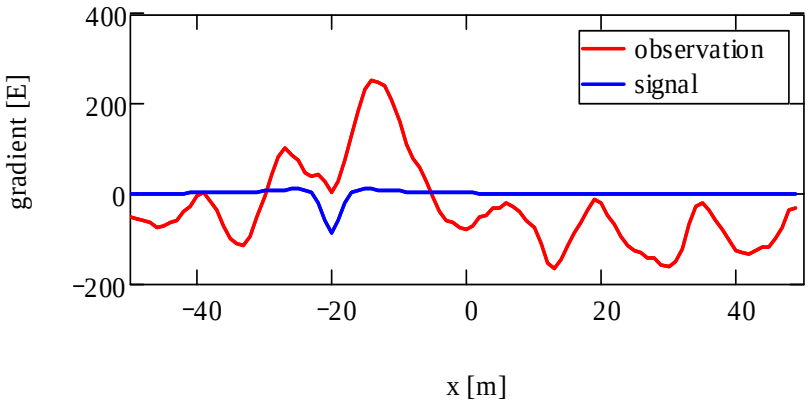


Figure 2: Observed Γ_{33} gradient (including a background gravitational field, instrument noise, and a signal), and the signal itself.

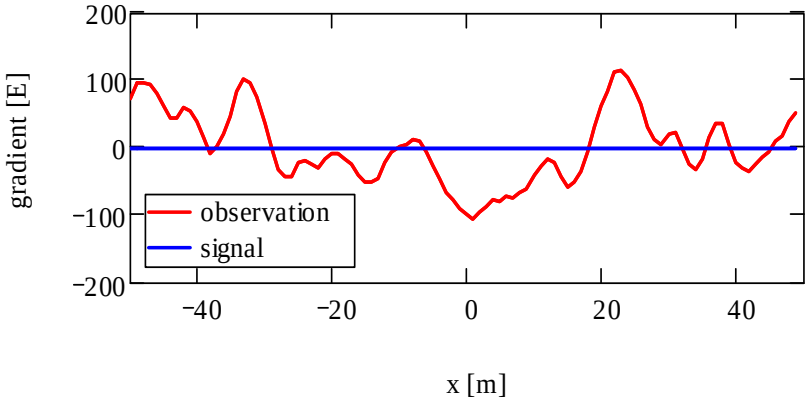


Figure 3: Observed $2\Gamma_{12}$ gradient (including a background gravitational field, instrument noise, and a signal), and the signal itself.

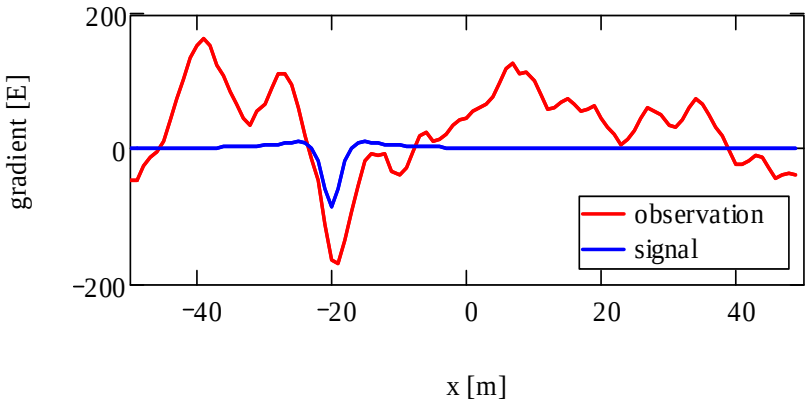


Figure 4: Observed $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$ gradient (including a background gravitational field, instrument noise, and a signal), and the signal itself.

Figures 5 and 6 show the matched filter output. The covariance function, ϕ_g, used in the algorithm corresponds exactly to the psd model that was used to generate the background gravitational gradients. A definite maximum in the output occurs at the true location of the anomaly in each case. The statistical tests offer a slightly less definite answer, as seen below.

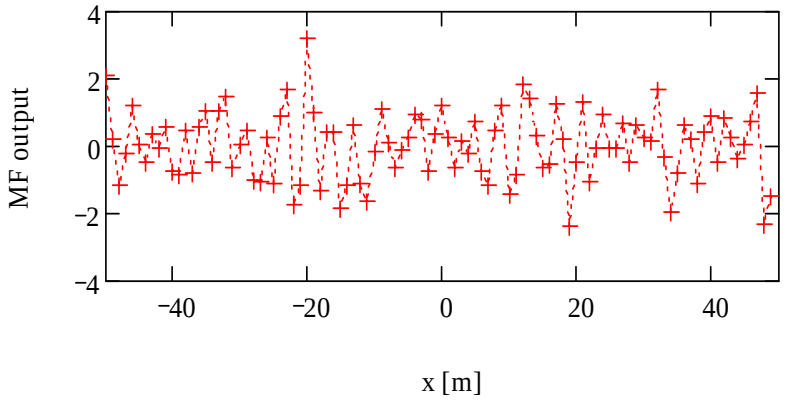


Figure 5: Output of matched filter applied to Γ_{33} observations.

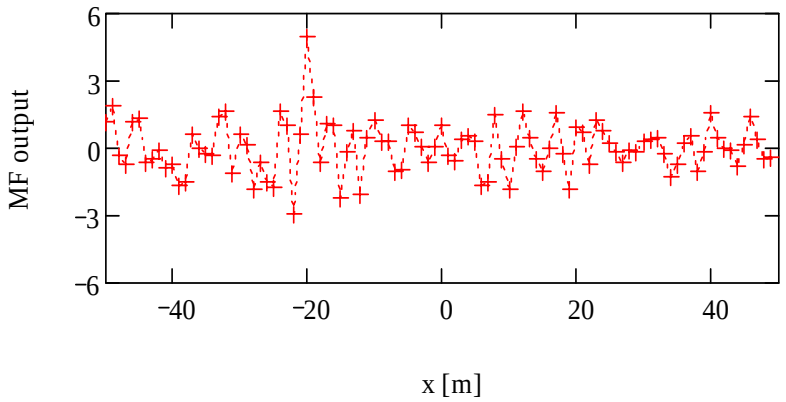


Figure 6: Output of matched filter applied to observations of $2\Gamma_{12}$ and $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$.

Table 1 summarizes the numerical results of the matched filter and the corresponding probabilities. All tests are performed with a significance level of $\alpha^{(A)} = \alpha^{(B)} = 0.05$. In the case of Γ_{33} observations, the null hypothesis in Setup A is not rejected, since $\bar{y} < \psi_0^{(A)}$, and we accept that there is no (detectable) signal. But, the probability of a miss is not close to zero (0.37) (i.e., there is a significant chance that we missed the signal, having accepted the null hypothesis). For the other set of observations, we reject the null hypothesis under Setup A since $\bar{y} > \psi_0^{(A)}$ and the probability of a miss is only 0.05. Under Setup B, the null hypothesis (that there is a detectable signal) is accepted in both cases; however, the probability of a false alarm is large in the first case (observing Γ_{33}) and only 0.05 in the other case.

We see that under Setup A, the statistical analysis apparently is more robust than under Setup B. Indeed, in the case of Γ_{33} observations, the power of the test under Setup A is $p = 1 - \beta = 0.73$ and under Setup B it is only $p = 1 - \beta = 0.09$. On the other hand, it is noted that in this particular case the test under Setup B gives the correct result, whereas with Setup A it does not. One must interpret the probabilities as just that – frequencies of occurrence of events realized from many experiments. When observing $2\Gamma_{12}$ and $\Gamma_{22} - \Gamma_{11}$, the power of the test is $p = 0.95$ under both setups, ostensibly because the signal-to-noise ratio is higher in this case and both tests readily validate our selection with a high probability value.

Table 1: Probabilities of Type II errors for MF example.

Observations	$\lambda^2 = SNR$	$\bar{y}$	$\arg\left(\max_x\left(y(x)\right)\right)$	$\psi_0^{(A)}$	$\psi_0^{(B)}$	$\beta^{(A)}$	$\beta^{(B)}$
Γ_{33}	13.158	3.195	$x = -20$ m	3.283	1.983	0.37	0.91
$2\Gamma_{12}$, $\Gamma_{22} - \Gamma_{11}$	18.020	4.910	$x = -20$ m	3.283	3.265	0.05	0.05

The next numerical example illustrates the results of the 2-D matched filter. The same background gradient field was used, but now there are three ($N_2 = 3$) parallel tracks of $N_1 = 99$ observation points each. Also, the same two gradients, $2\Gamma_{12}$ and $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$, as in Section VI were observed (i.e., simulated) at each point. The along-track spacing is the same as before, $\Delta x_1 = 1$ m, and the tracks are spaced at $\Delta x_2 = 3$ m. Also, the instrument noise is white with standard deviation, $\sigma = 3$ E. The anomaly is as before, 1 m $\times$2 m in cross-section, 100 m long, perpendicular to the tracks, and at a depth (to the top) of $d = 2$ m. The geometric center of the anomaly is again located at $x_1 = -20$ m, $x_2 = 0$ m, $x_3 = -3$ m. Figure 7 shows the profiles of the observations, z_{j_1, j_2}.

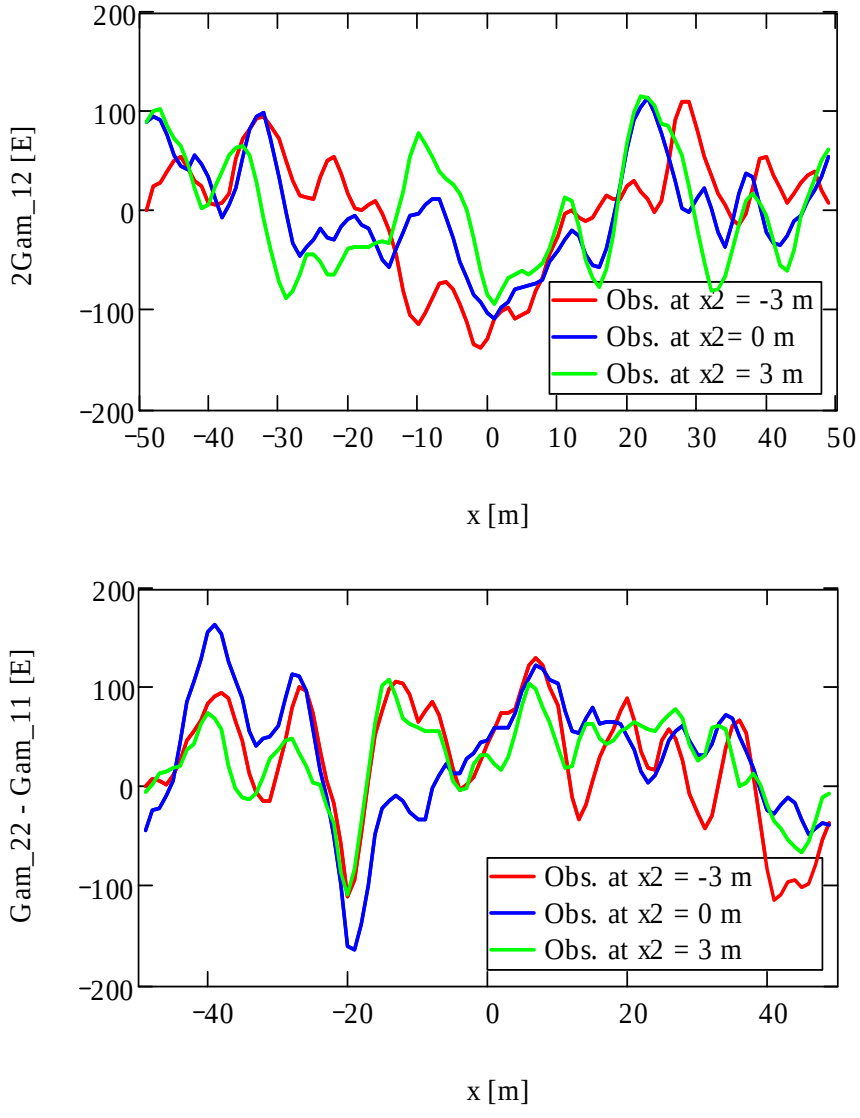


Figure 7: Profiles of $2\Gamma_{12}$ (top) and $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$ (bottom) observations (including background gravitational field, instrument noise and anomaly signal).

The matched filter output is shown in Figure 8, where the outputs for the three tracks are offset intentionally, otherwise they would essentially overlap. The signal-to-noise ratio in this case is $\lambda^2 = 36.58$ and the maximum filter outputs per track are

$$\max_{j_1}\left(y_{j_1,-1}\right) = 6.87, \quad \max_{j_1}\left(y_{j_1,0}\right) = 6.92, \quad \max_{j_1}\left(y_{j_1,1}\right) = 6.64, \tag{76}$$

precisely at the location of the anomaly. We see that the filter output is highly correlated in the cross-track direction, which would occur even in the absence of the anomaly embedded in the noise (see the next section). Therefore, since the pdf (equation (68)) of the maximum output with no signal requires all outputs to be independently distributed, the determination of probabilities of a miss or a false alarm cannot be computed as it was in the case of a single track.

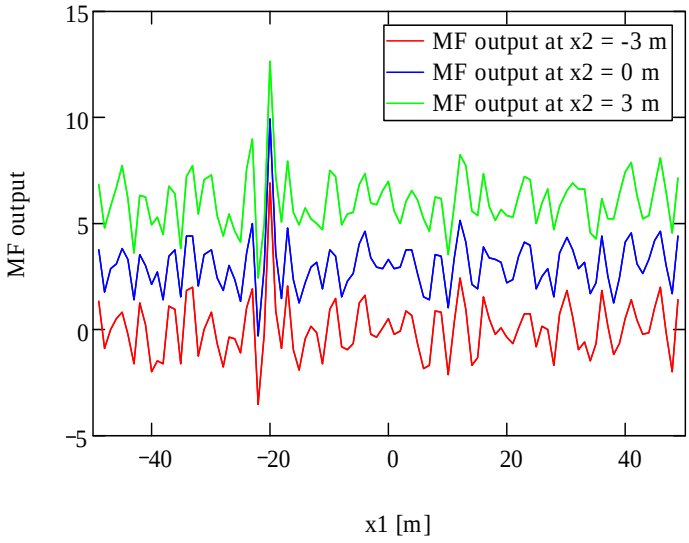


Figure 8: Output of the 2-D matched filter applied to observations of $2\Gamma_{12}$ and $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$ on three parallel tracks. The outputs for $x_2 = 0$ m and $x_2 = 3$ m are offset intentionally for clarity.

VIII Correlation of the MF Output

In order to understand the correlative nature of the filter output for multiple tracks in the numerical example of the previous section, we consider the case when the noise is uncorrelated ($\phi = I$) and there is no signal in the observations ($z = n$). For multiple tracks, the output becomes (equations (44) and (45))

$$y_{r_1,r_2} = \frac{1}{\lambda} \sum_{j_1=-(N_1-1)/2}^{(N_1-1)/2} \sum_{j_2=-(N_2-1)/2}^{(N_2-1)/2} \overline{s}_{-j_1,-j_2} \tilde{n}_{r_1-j_1,r_2-j_2} \, . \tag{77}$$

Let us further simplify this case by assuming that the putative signal is an along-track impulse:

$$s_{j_1,j_2} = \begin{cases} 1, & j_1 = 0 \\ 0, & j_1 \neq 0 \end{cases} \tag{78}$$

for all j_2. That is, the signal is constant in the cross-track direction. The signal-to-noise ratio is then $\lambda^2 = \overline{\mathbf{s}}_-^{\mathrm{T}} \phi_-^{-1} \overline{\mathbf{s}}_- = N_2$; and the filter output becomes

$$y_{r_1,r_2} = \frac{1}{\sqrt{N_2}} \sum_{j_2=-(N_2-1)/2}^{(N_2-1)/2} \tilde{n}_{r_1,r_2-j_2} \,. \tag{79}$$

The difference in the output between adjacent tracks is

$$\begin{aligned} \sqrt{N_2}\left(y_{r_1,r_2} - y_{r_1,r_2-1}\right) &= \sum_{j_2=-(N_2-1)/2}^{(N_2-1)/2} \tilde{n}_{r_1,r_2-j_2} - \sum_{j_2=-(N_2-1)/2}^{(N_2-1)/2} \tilde{n}_{r_1,r_2-1-j_2} \\ &= \sum_{j_2=-(N_2-1)/2}^{(N_2-1)/2} \tilde{n}_{r_1,r_2-j_2} - \sum_{j_2=-(N_2-1)/2+1}^{(N_2-1)/2+1} \tilde{n}_{r_1,r_2-1-(j_2-1)} \\ &= \sum_{j_2=-(N_2-1)/2}^{(N_2-1)/2} \tilde{n}_{r_1,r_2-j_2} - \sum_{j_2=-(N_2-1)/2+1}^{(N_2-1)/2+1} \tilde{n}_{r_1,r_2-j_2} \\ &= \tilde{n}_{r_1,r_2+(N_2-1)/2} - \tilde{n}_{r_1,r_2-(N_2-1)/2-1} \end{aligned} \tag{80}$$

From equation (56), $y_{r_1,r_2} - y_{r_1,r_2-1} = 0$, on account of the periodic extension assumed for the background noise. That is, in this simple case, the filter output is identical for all tracks. Figure 9 shows an example where the noise alternatively is extended by zeros (analogous to equation (7)). There is still significant correlation between the tracks, because only one of the summands in equation (79) is different from one track to its adjacent track. This is due to the linear feature of the putative signal that is perpendicular to all tracks.

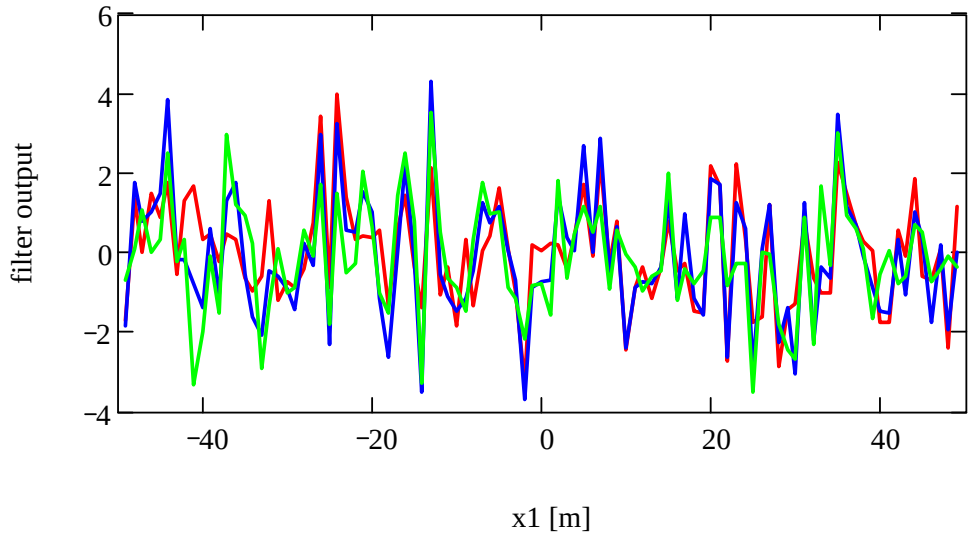


Figure 9: Filter output along three parallel tracks (three colors) for the case of unit-variance white noise, a delta function for the putative signal, no signal in the observations, and the extension of observations with zeros.

Although we used a very simple example here to illustrate the cross-track correlation of the filter output, the result easily is seen to extend to the more general case. We may conclude that the assumed signal in the design of the filter function is what causes the correlation of the output among the tracks. It means also that the statistical analysis of Section VI is not directly applicable for multiple tracks, and further studies are needed to develop it for this case. One option is to conduct the matched filter and corresponding statistical analysis separately for each track.

IX Arbitrary Orientation of the Signal

As a further numerical example of the matched filter (MF) for multiple tracks of data, consider the case of an arbitrarily oriented anomaly. As before, the same background gradient field is used with three ($N_2 = 3$) parallel tracks of $N_1 = 99$ observation points each. Also, the same two gradients, $2\Gamma_{12}$ and $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$, are observed at each point. The only change is the assumed orientation of the anomaly in the filter function. The along-track and cross-track spacings are the same as before: $\Delta x_1 = 1$ m, $\Delta x_2 = 3$ m. Also, the instrument noise is assumed to be white with standard deviation, $\sigma = 3$ E. The anomaly is a right rectangular prism, as before, but at arbitrary (presumed known) orientation with respect to the tracks. The other parameters of the prism, (width, height, length, and depth) are also assumed known. The true location of the anomaly is such that its geometric center has coordinates: $x_1 = -20$ m, $x_2 = 0$ m, $x_3 = -3$ m. Its true dimensions are 1 m $\times$ 2 m in cross-section and 100 m in length. The true orientation is 40° in azimuth (zero degrees for the dip and twist angles).

Figure 10 shows the plan view and profiles of the signals along the tracks. Figure 11 shows the observations, z_{j_1,j_2}, which include the background, the signal of the anomaly, and the

simulated instrument noise. The results of the matched filter are shown in Figure 12 for differently assumed azimuths of the anomaly with additional information provided in Table 2.

We see that only by assuming the correct azimuth, $\alpha = 40°$, does the filter yield the correct location of the anomaly. However, only the middle track is able to detect this correct location (the maximum output occurs at $x_1 = -20$ m). The two-dimensional filter is a convolution in two dimensions, and we must assume that the observations are extended beyond the observation grid in some way. We chose a periodic extension, which for the non-perpendicularly oriented anomaly yields a significant distortion in the signal in the cross-track direction. The situation does not improve if the observations are extended by zeros. This is believed to be the reason that the filter is unable to detect the rather strong signal on the outer tracks (recall that when the anomaly is perpendicular to the tracks, the output on all tracks easily identifies the location, if the signal is strong, as in this case; see Figure 8).

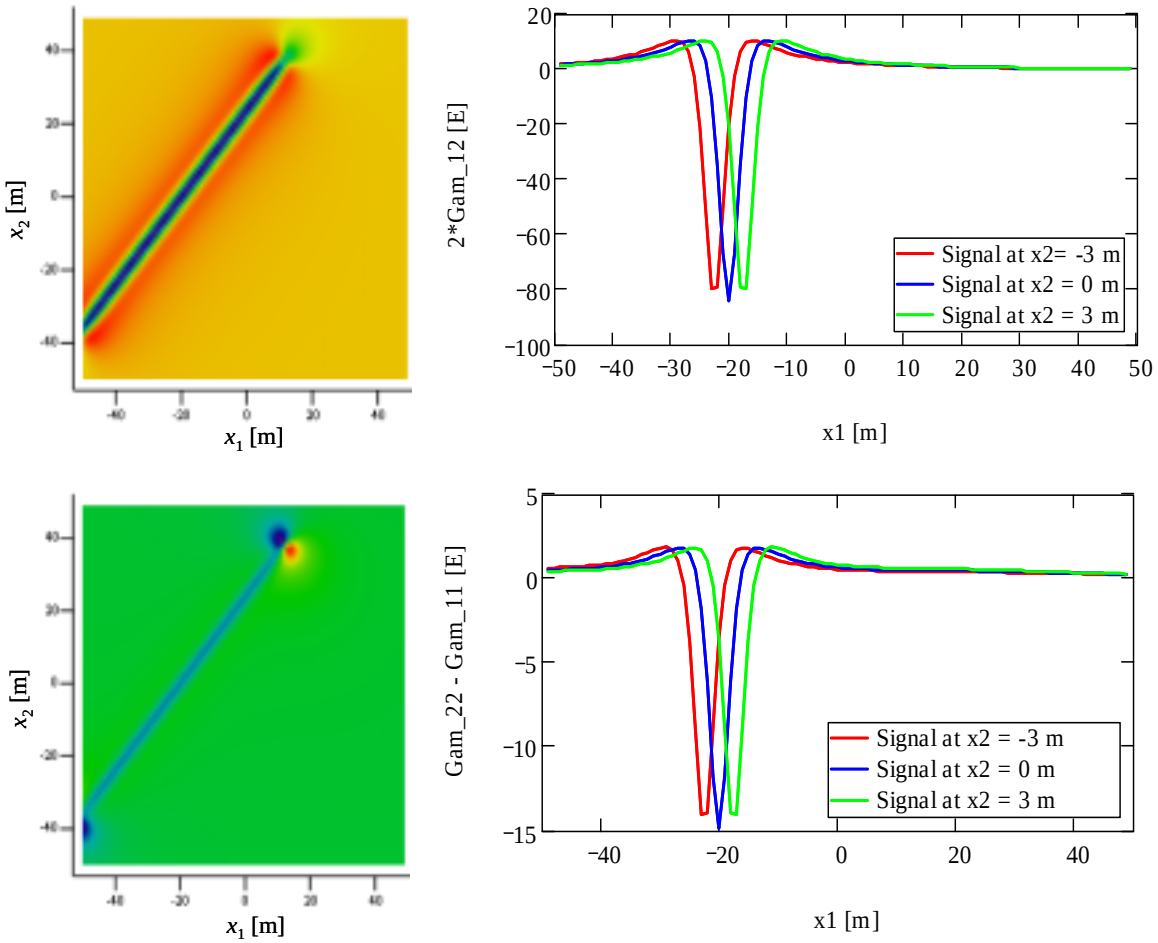


Figure 10: Plan view of $2\Gamma_{12}$ (top) and of $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$ (bottom) due to the anomaly signal and corresponding profiles along three parallel tracks separated by $\Delta x_2 = 3$ m.

We conclude that the matched filter in two dimensions is not particularly useful for this application. Instead, given multiple tracks of data crossing a linear anomaly, each track of data

should be processed separately by the one-dimensional matched filter. This also has the advantage that the output is amenable to a probabilistic interpretation, where the probability of a miss or a false alarm can be calculated on the basis of known probability distributions.

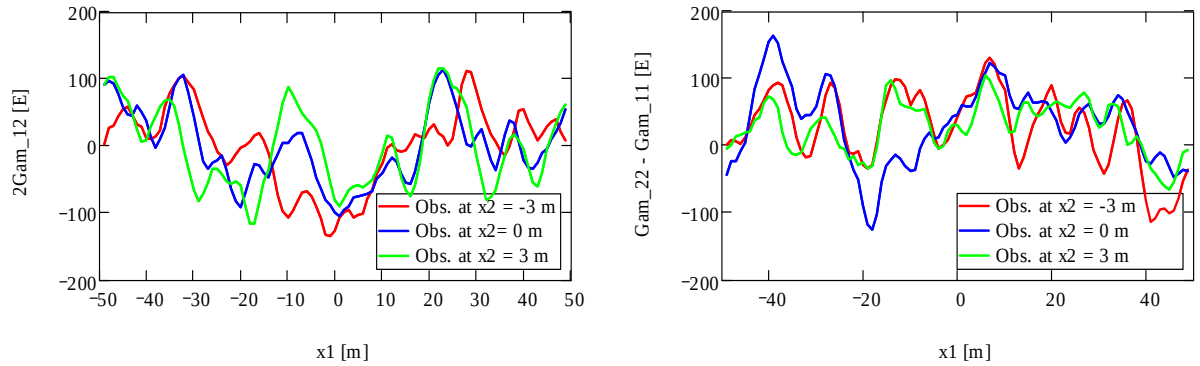


Figure 11: Profiles of $2\Gamma_{12}$ (left) and $\Delta\Gamma = \Gamma_{22} - \Gamma_{11}$ (right) observations (including background gravitational field, instrument noise and anomaly signal).

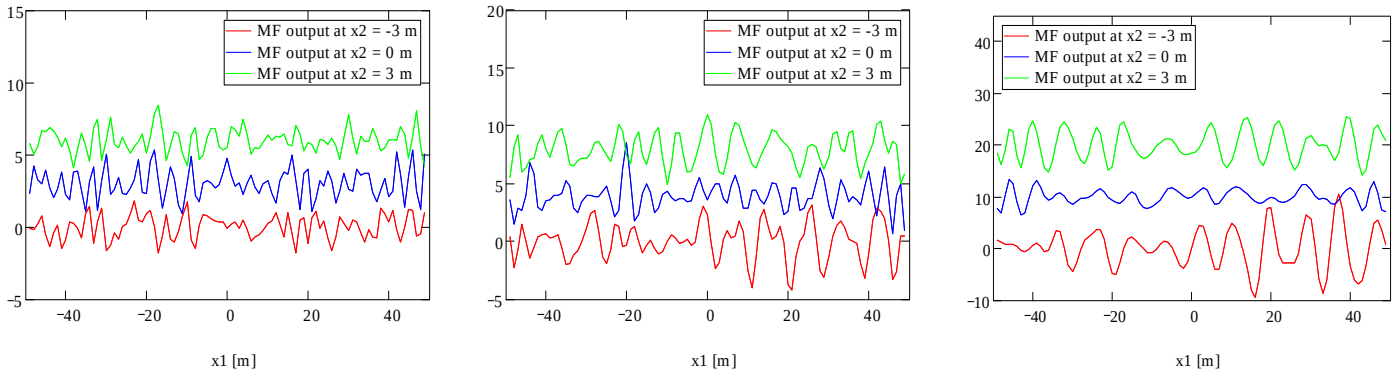


Figure 12: MF outputs along three parallel tracks (three colored lines), assuming a particular azimuthal orientation of the anomaly: 20° (left), 40° (middle) and 60° (right).

Table 2: Results of the 2-D matched filter for different assumed azimuth of the anomaly.

assumed azimuth of anomaly	signal-to-noise, λ^2	location of max output, $x_{1,m}$, per track		
		$x_2 = -3$ m	$x_2 = 0$ m	$x_2 = 3$ m
$\alpha = 20°$	36.9	48 m	-18 m	30 m
$\alpha = 40°$	30.9	42 m	**-20 m**	16 m
$\alpha = 60°$	28.0	37 m	43 m	30 m

Figure 13 shows the MF output if each track is treated individually. Table 3 shows additional information on the matched filter results. For each track, the correct location of the anomaly is identified consistently if the correct azimuth is chosen in the filter (note that the

anomaly crosses the $x_2 = -3$ m track at (approximately) $x_1 = -23$ m and the $x_2 = 3$ m track at $x_1 = -17$ m). In two cases the correct location is also identified if the incorrect azimuth ($\alpha = 20°$) is built into the filter. Considering the results for all three tracks, we may conjecture that the anomaly, if correctly detected, is oriented at an azimuth of $\alpha = 40°$. An analysis of the calculated probabilities follows.

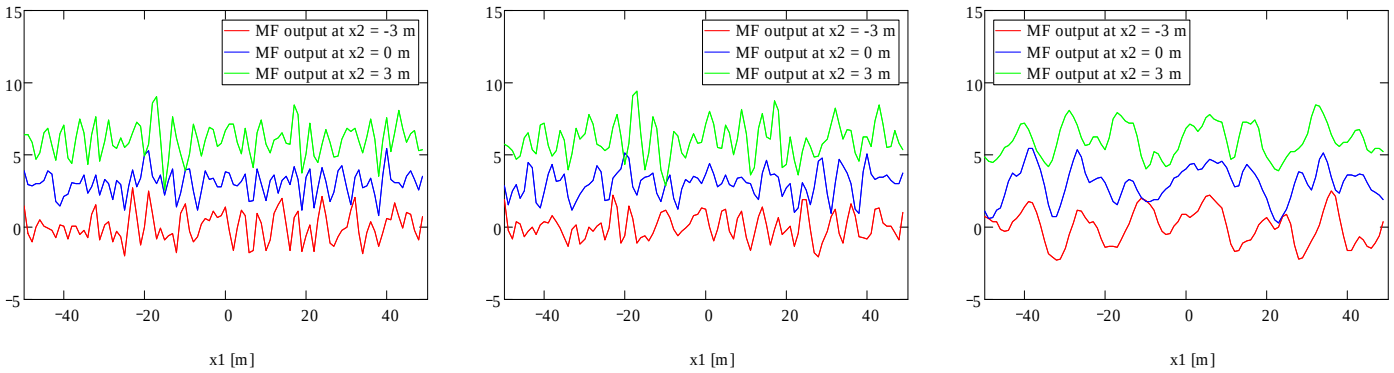


Figure 13: MF outputs along three parallel tracks (three colored lines), assuming a particular azimuthal orientation of the anomaly: 20° (left), 40° (middle) and 60° (right). Each track of data was filtered individually.

Table 3: Results of the 1-D matched filter for different assumed azimuths of the anomaly.

assumed azimuth of anomaly	signal-to-noise, λ^2	location of max output, $x_{1,m}$, per track		
		$x_2 = -3$ m	$x_2 = 0$ m	$x_2 = 3$ m
$\alpha = 20°$	17.2	**-23 m**	40 m	**-17 m**
$\alpha = 40°$	9.7	**-23 m**	**-20 m**	**-17 m**
$\alpha = 60°$	3.9	36 m	-38 m	32 m

Tables 4 and 5 list the results of hypothesis testing under the two Setups A and B, respectively. All tests are performed with a significance level of $\alpha^{(A)} = \alpha^{(B)} = 0.05$. Under Setup A, the threshold is $\psi_0^{(A)} = 3.283$, and the null hypothesis (there is no anomaly) is accepted in every case except for the $x_2 = 3$ m track when the azimuth was correctly modeled. Thus, we get the correct outcome on only one out of three tracks when the orientation is correctly modeled. Does the probability of a miss give a reasonable characterization of these actions?

The computed probabilities of a miss are small for the incorrect azimuth model, $\alpha = 20°$, which is consistent, in principle, with our acceptance of the null hypothesis in these cases. That is, the orientation is incorrect (so our action is correct and we do not miss the correctly oriented anomaly); but, in fact, we actually do locate the anomaly in two out of three tracks, so one might argue that the *POM* should be higher.

For the other incorrect model, $\alpha = 60°$, the probability of a miss is very high, which in this case definitely is inconsistent with the result from the hypothesis test (we accept the null

hypothesis and it is the correct thing to do – the probability of a miss should be low). For the correctly modeled orientation ($\alpha = 40°$), the calculated probability of a miss is about even (we do indeed miss the anomaly on two tracks, by accepting the null hypothesis, but find the anomaly on the third track by rejecting the null hypothesis). So, in this case the probability is consistent with the outcome.

Under Setup B, the null hypothesis (there is an anomaly) is accepted in all but one case. This is certainly the correct action for the orientation model, $\alpha = 40°$ (that is, we do find the anomaly on all three tracks). However, the calculated probability of a false alarm is very high (1.00) in this case, which is not consistent with the fact that we correctly locate the anomaly – it was not a false alarm. When the incorrect orientation model is $\alpha = 60°$, we accept the hypothesis that the anomaly is detected, but clearly we do not locate it correctly, and the high *POF* (1.00) is consistent with that action. On the other hand, when the incorrect orientation model is $\alpha = 20°$, the *POF* is neither high nor low (0.64), but, in principle, we should consistently reject the null hypothesis (that the correctly modeled anomaly is detected) and the *POF* should be low. In fact, we do locate the anomaly consistently with our acceptance of the null hypothesis, but under false pretenses (i.e., the model is incorrect).

Table 4: Accept $H_0^{(A)}$ (no anomaly) if $\bar{y} < \psi_0^{(A)}$. Bolded output rejects $H_0^{(A)}$.

assumed azimuth of anomaly	$x_2 = -3$ m			$x_2 = 0$ m			$x_2 = 3$ m		
	$\bar{y}$	$\psi_0^{(A)}$	$\beta^{(A)}$	$\bar{y}$	$\psi_0^{(A)}$	$\beta^{(A)}$	$\bar{y}$	$\psi_0^{(A)}$	$\beta^{(A)}$
$\alpha = 20°$	2.70	3.283	0.19	2.43	3.283	0.19	3.00	3.283	0.19
$\alpha = 40°$	2.16		0.57	2.10		0.57	**3.35**		0.57
$\alpha = 60°$	2.48		0.91	2.43		0.91	2.45		0.91

Table 5: Accept $H_0^{(B)}$ (anomaly) if $\bar{y} > \psi_0^{(B)}$. Bolded output rejects $H_0^{(B)}$.

assumed azimuth of anomaly	$x_2 = -3$ m			$x_2 = 0$ m			$x_2 = 3$ m		
	$\bar{y}$	$\psi_0^{(B)}$	$\beta^{(B)}$	$\bar{y}$	$\psi_0^{(B)}$	$\beta^{(B)}$	$\bar{y}$	$\psi_0^{(B)}$	$\beta^{(B)}$
$\alpha = 20°$	2.70	2.50	0.64	**2.43**	2.50	0.64	3.00	2.50	0.64
$\alpha = 40°$	2.16	1.47	1.00	2.10	1.47	1.00	3.35	1.47	1.00
$\alpha = 60°$	2.48	0.33	1.00	2.43	0.33	1.00	2.45	0.33	1.00

From these simple tests, the probabilities under Setup A appear to be more robust (from the viewpoint of being more informative about our actions based on the hypothesis tests) than under Setup B when the correct orientation model is used, which is consistent with more extensive tests based on Monte Carlo simulations discussed in the next section. However, for the incorrectly modeled orientations, the probabilities under either Setup were not uniformly informative. That is, neither the computed *POM* nor the computed *POF* necessarily provided a reasonable characterization of our action based on the hypothesis test. In fact, in some cases (for the incorrectly modeled orientation), the computed probability characterized our action as just the opposite of what it should have done. For example, when we correctly accepted that the putative

anomaly does not exist, we obtained a large *POM* (last row of Table 4). Clearly, the computation of probabilities for the output of an incorrectly modeled matched filter requires further study and, indeed, one should perhaps not expect these probabilities to be the same as for the correctly modeled MF.

X Validation of Statistical Performance

In developing the statistical tests for the MF with correlated noise, a number of assumptions were made regarding the probability density of the filter output. These densities assumed either the strong presence or complete absence of a signal, as well as the independence among the MF outputs in the latter case. The following simulations illustrate the performance of the tests under these assumptions and for signals ranging from strong to nonexistent.

We continue with the example of detecting a simple sub-surface structure from measurements of gravity gradients on the surface along a profile that intersects the signal at right angles. In order to study the detectability of a specific mass density anomaly in a statistical setting, both the background gravity field and the instrument noise were determined in Monte Carlo fashion from simulated random processes. As before, the gradiometer *measurement* error was assumed to be white and Gaussian with zero mean and a given standard deviation, $\sigma_\varepsilon = 3$ E. The background gradient field was also assumed to be a known stationary, Gaussian, stochastic process. Moreover, the component covariance functions were used to generate realizations of each process (Appendix E).

We considered two sets of gradients, either $\{\Gamma_{33}\}$ or $\{\Gamma_{13}, \Gamma_{33}\}$, observed at regularly spaced points along the track, with $\Delta x_1 = 1$ m. No additional gradients were tested since for this signal the gravity gradients, Γ_{12}, Γ_{22}, and Γ_{23}, are zero; and by Laplace's field equation, $\Gamma_{11} = -\Gamma_{33}$. The observations of the vertical gradient, Γ_{33}, (and analogously for the cross-gradient, Γ_{13}) were simulated according to

$$z_j = \Gamma_{33}^s\left(\left(j-m\right)\Delta x_1\right) + \Gamma_{33}^n\left(j\Delta x_1\right) + \varepsilon_j, \quad j = 1,\ldots,100, \tag{81}$$

where Γ_{33}^s is the gradient of the anomaly, $m = 30$, Γ_{33}^n is a realization of the background gradient, and ε_j is a realization of the measurement noise. As another example, analogous to Figures 2 and 5, Figure 14 (top) shows the observations, equation (81), where the top of the anomaly is at 2 m depth. While the signal of the anomaly is barely noticeable by visual inspection of this single profile, the maximum filter output (Figure 14, bottom) correctly identifies its location. It is also noted from Figure 14 that the observations come from a decidedly correlated process, but the MF output is much less correlated (except near the signal), which supports our assumption on the probability density of the output in the absence of a signal.

Among 1000 such realizations of the background field (and measurement noise), the matched filter correctly detected the signal of the density anomaly most of the time when its depth was just 1 m or 2 m. An indication within 1 m of the correct horizontal location was considered a success. At 3 m depth the maximum filter output still was situated more often at the correct location than at any other particular point. When the signal was deeper, or did not exist, the maximum filter output fairly uniformly mis-located it. The percentages of successfully identifying the anomaly based on the hypothesis tests, together with the percentage errors in

those choices are summarized in Table 6 (Setup A) and Table 7 (Setup B). These results depend, though not with high sensitivity, on the thresholds computed under each of the two Setups. The Tables include cases of deliberate mismodeling of the filter function with respect to the assumed depth of the signal ($\overline{\boldsymbol{s}} \neq \boldsymbol{s}$ in equation (28)).

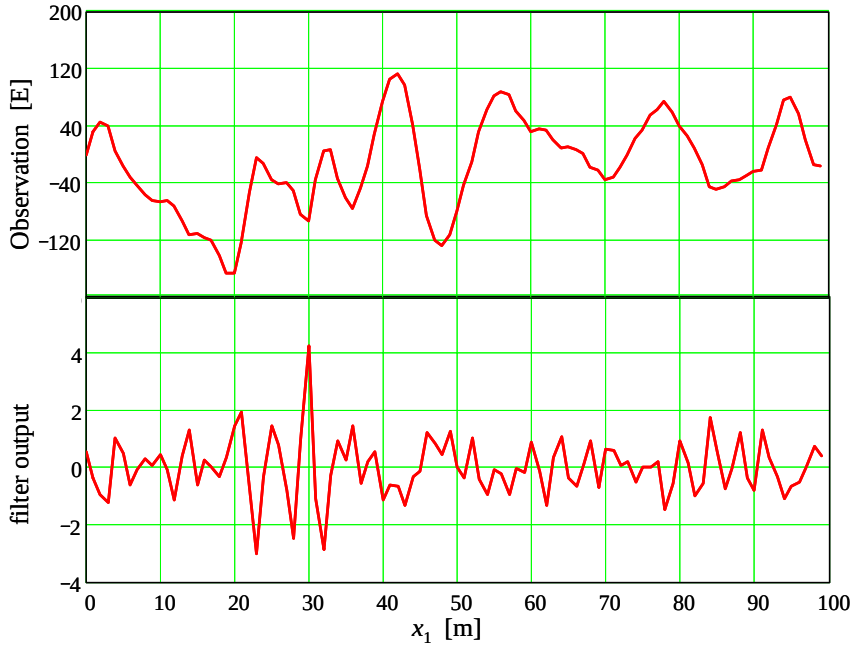


Figure 14: Top: Along-track observations of Γ_{33} including the signal from the mass anomaly at $x_1 = 30$ m and 2 m depth, the background gradient field (see Figure E.1), and white measurement noise ($\sigma_I = 3$ E). Bottom: matched-filter output, with its maximum indicating the location of the anomaly.

Table 6: Percentages of declaring existence or non-existence of the signal; with parenthetical percentages of corresponding errors in the declaration (Setup A).

true depth	= 1 m	= 2 m	= 3 m	= 4 m	= 1 m	= 2 m	= 3 m	= 4 m
filter depth	Declare Signal (reject $H_0^{(A)}$)				Declare No Signal (accept $H_0^{(A)}$)			
	$\{\Gamma_{33}\}$							
1 m	**100 (0.0)**	47.4 (9.3)	6.1 (93.4)	5.9 (98.3)	**0.0 (-)**	52.6 (54.0)	93.9 (6.4)	94.1 (4.5)
2 m	100 (0.0)	**64.0 (4.7)**	6.1 (85.2)	5.2 (98.1)	0.0 (-)	**36.0 (63.9)**	93.9 (9.3)	94.8 (5.2)
3 m	100 (0.0)	43.0 (9.1)	**6.0 (78.3)**	5.2 (90.4)	0.0 (-)	57.0 (55.8)	**94.0 (12.6)**	94.8 (5.3)
4 m	100 (0.0)	15.7 (21.0)	5.0 (74.0)	**4.5 (93.3)**	0.0 (-)	84.3 (37.2)	95.0 (13.6)	**95.5 (7.0)**
	$\{\Gamma_{13}, \Gamma_{33}\}$							
1 m	**100 (0.0)**	66.8 (3.8)	6.6 (90.9)	6.2 (96.8)	**0.0 (-)**	31.2 (66.7)	93.4 (7.4)	93.8 (4.8)
2 m	100 (0.0)	**85.1 (1.1)**	6.2 (74.2)	5.0 (92.0)	0.0 (-)	**14.9 (68.5)**	93.8 (13.2)	95.0 (5.5)
3 m	100 (0.0)	63.4 (3.2)	**7.6 (53.9)**	5.0 (84.0)	0.0 (-)	36.6 (65.8)	**92.4 (19.3)**	95.0 (8.4)
4 m	99.9 (0.0)	24.0 (7.5)	4.6 (47.8)	**3.1 (74.2)**	0.1 (0.0)	76.0 (48.7)	95.4 (19.4)	**96.9 (9.1)**

Under Setup A, the null hypothesis (no signal) is selected more often as the depth of the anomaly increases and consequently as the signal is more thoroughly buried in the noise (right half of Table 6). The percentage of the error in this choice also decreases correspondingly. This percent error (in parentheses) is determined from the ratio of errors made in this decision relative to all decisions to accept the null hypothesis. This is not the probability of a miss as defined by equation (74) where the sample space comprises all cases of actual detection, i.e., $H_1^{(A)}$ is true (see also equation (83) and the corresponding discussion, below). Conversely, the rate of rejecting the null hypothesis decreases with increasing anomaly depth, and the percent error in this choice increases (left half of Table 6).

With an incorrectly modeled anomaly depth, the correct hypothesis is somewhat less often selected when the anomaly is at a depth of 2 m. Otherwise, the modeled depth (if within a few meters of the true depth) appears to have little impact on the results.

Table 7: Percentages of declaring existence or non-existence of the signal; with parenthetical percentages of corresponding errors in the declaration (Setup B).

true depth	= 1 m	= 2 m	= 3 m	= 4 m	= 1 m	= 2 m	= 3 m	= 4 m
filter depth	Declare Signal (accept $H_0^{(B)}$)				Declare No Signal (reject $H_0^{(B)}$)			
	$\{\Gamma_{33}\}$							
1 m	**94.5 (0.0)**	0.0 (-)	0.0 (-)	0.0 (-)	**5.5 (100)**	100 (71.4)	100 (6.4)	100 (4.3)
2 m	100 (0.0)	**99.7 (15.9)**	92.0 (90.7)	91.3 (95.4)	0.0 (-)	**0.3 (66.7)**	8.0 (12.5)	8.7 (9.2)
3 m	100 (0.0)	100 (29.1)	**100 (86.9)**	100 (94.5)	0.0 (-)	0.0 (-)	**0.0 (-)**	0.0 (-)
4 m	100 (0.0)	100 (56.2)	100 (85.8)	**100 (93.0)**	0.0 (-)	0.0 (-)	0.0 (-)	**0.0 (-)**
	$\{\Gamma_{13}, \Gamma_{33}\}$							
1 m	**94.7 (0.0)**	0.0 (-)	0.0 (-)	0.0 (-)	**5.3 (100)**	(100) (87.0)	100 (7.5)	100 (4.7)
2 m	100 (0.0)	**96.9 (3.8)**	35.2 (81.8)	32.3 (93.8)	0.0 (-)	**3.1(38.7)**	64.8 (11.7)	67.7 (5.3)
3 m	100 (0.0)	100 (14.5)	**100 (78.7)**	100 (91.2)	0.0 (-)	0.0 (-)	**0.0 (-)**	0.0 (-)
4 m	100 (0.1)	100 (40.8)	100 (79.1)	**100 (90.4)**	0.0 (-)	0.0(-)	0.0 (-)	**0.0 (-)**

For Setup B, the null hypothesis (the signal exists) is almost always chosen, whether the signal is truly detected or not, and even if the anomaly depth is modeled incorrectly (left half of Table 7). The only exception occurs when the anomaly depth is mis-modeled at 1 m (or also 2 m for the two-gradient observations). The a posteriori rate of wrongly choosing the null hypothesis increases as the signal vanishes in the noise. Clearly, the results of these simulations (Tables 6 and 7) show that the correct decision is made more often if the hypothesis and corresponding threshold are based on Setup A. On the other hand, the indiscriminate declaration of the existence of the signal under Setup B is compensated, as shown below, by an appropriate predicted probability of a false alarm.

The performance of the predicted (theoretical) probabilities of a false alarm or a miss depends on the validity of the underlying statistics of the matched filter output. Table 8 shows for the gradient, Γ_{33}, that the *empirical* mean and standard deviation of the maximum output, $\bar{y}$, determined from the 1000 simulations, correspond less to the theoretical values implied by the Gaussian distribution, as the signal becomes less detectable (depth increases) and the distribution of the *maximum* output changes to that of an order-statistic. When the observations contain no signal ($\boldsymbol{s} = \mathbf{0}$), the theoretical and empirical statistics of the filter output, y_r, agree more closely, as Table 9 also shows. The empirical standard deviation in this case was solved as the parameter in a least-squares fit of equation (68) to a histogram of $\bar{y}$. The empirical mean of y_r in all cases was numerically indistinguishable from its theoretical value of zero.

Table 8: Empirical statistics of the filter output and its maximum, determined from 1000 simulations of the background gravity field (Γ_{33}). The theoretical mean, λ, is computed according to equation (29). The theoretical standard deviation in all cases is unity.

depth to top of assumed signal, $\bar{s}$	$s = \bar{s}$			$s = 0$
	emp. mean of $\bar{y}$	theor. mean of $\bar{y}$	emp. st.dev. of $\bar{y}$	emp. st.dev. of y
1 m	22.968	22.996	1.001	0.995
2 m	3.696	3.627	0.862	1.005
3 m	2.528	1.055	0.466	0.993
4 m	2.383	0.517	0.484	0.936

Errors in the presumed statistics, mean and standard deviation, of the filter output affect the determination of the threshold, ψ_0, to which, however, the acceptance or rejection of hypotheses under the two Setups is not particularly sensitive. The potentially greater effect of such errors is on the probability density function of the maximum filter output, which, in turn, affects the value of the computed probability of a Type-II error (β). We compared the values of β computed from theoretical and empirical statistics of the filter output. For Setup A, the *theoretical* threshold is determined numerically from equation (71), based on the theoretical statistics for the MF output when there is no signal (zero mean and unit variance), and with $N = 100$ since the maximum output is determined from 100 samples. The probability of a Type I error was chosen to be $\alpha^{(A)} = 0.05$. The *empirical* threshold is computed by replacing in equation (71) the theoretical unit standard deviation of the filter output with the empirical value shown in Table 8. The Type-II error probability, the *POM* in this case, is computed either by equation (74) using theoretical statistics and threshold, or using the empirical values by

$$\beta^{(A)} = \frac{1}{2} + \frac{1}{2}\,\mathrm{erf}\left(\frac{\left(\psi_0^{(A)}\right)_{\mathrm{emp}} - \mu_{\mathrm{emp}}^{\bar{y}}}{\sqrt{2}\,\sigma_{\mathrm{emp}}^{\bar{y}}} \right), \tag{82}$$

where $\mu_{\mathrm{emp}}^{\bar{y}}$ and $\sigma_{\mathrm{emp}}^{\bar{y}}$ are the empirical mean and standard deviation of the maximum output (Table 8).

Table 9 shows that $\beta^{(A)}$ in several cases is slightly (though perhaps not significantly) better predicted using the empirical MF output statistics. That is, column 6 of this Table, showing the still theoretically formulated *POM*, compares better than column 4 against the last column of Table 9 that shows the empirically determined *POM*. The empirical *POM* is computed on the basis of the 1000 simulated background noise fields by emulating equation (74). This formula gives the frequency of instances when the hypothesis test failed (the null hypothesis was wrongly accepted) among all instances that the signal is detectable ($H_1^{(A)}$ is true). Thus,

$$\text{emp. } POM = \frac{\#\text{ successes when } \bar{y} < \psi_0^{(A)}}{\#\text{ successes}}, \tag{83}$$

where a "success" means the filter correctly identifies the signal. The essential agreement between the empirical and theoretical *POM*, whether the latter is based on theoretical or empirical statistics, is useful since empirical statistics may not be obtained easily in practice.

Table 9: *POM*s determined either from theoretical or empirical thresholds for various depths of the signal. The significance level of the hypothesis test is $\alpha^{(A)} = 0.05$.

Observations	depth	theor. $\psi_0 \Rightarrow$ theor. POM		emp. $\psi_0 \Rightarrow$ theor. POM		***emp.* POM**
$\{\Gamma_{33}\}$	1 m	3.284	0.00	3.266	0.00	0.0
	2 m		0.37	3.299	0.32	0.27
	3 m		0.99	3.262	0.94	0.90
	4 m		1.00	3.073	0.92	0.96
$\{\Gamma_{13}, \Gamma_{33}\}$	1 m	3.284	0.00	3.269	0.00	0.0
	2 m		0.15	3.332	0.15	0.11
	3 m		0.97	3.231	0.92	0.84
	4 m		0.99	3.076	0.94	0.92
	5 m		1.00	2.829	0.90	0.97

Errors in the statistics of the output seem to affect more the *POF*, $\beta^{(B)}$, under Setup B (compare columns 4 and 6 in Table 10), although a useful quantitative assessment exists only for a single depth. Using theoretical parameters, $\beta^{(B)}$ was computed by equation (75). Empirical values were substituted as before to compute $\beta^{(B)}$ in column 6 of Table 10. The last column represents the empirical *POF*, computed from the 1000 realizations of the simulated background field and noise

$$\text{emp. } POF = \frac{\#\text{ failures when } \bar{y} > \psi_0^{(B)}}{\#\text{ failures}}, \tag{84}$$

where a "failure" means the filter does not correctly identify the signal. It is more difficult to claim from these simulations that the empirical *POF* agrees with the $\beta^{(B)}$ computed from either the empirical or theoretical statistics of the output.

Table 10: *POF*s determined either from theoretical or empirical thresholds for various depths of the signal. The significance level of the hypothesis test is $\alpha^{(B)} = 0.05$.

Observations	depth	theor. $\psi_0 \Rightarrow$ theor. POF		emp. $\psi_0 \Rightarrow$ theor. POF		***emp.* POF**
$\{\Gamma_{33}\}$	1 m	21.351	0.00	21.322	0.00	0.0
	2 m	1.982	0.91	2.278	0.69	0.99
	3 m	-0.588	1.00	1.761	0.98	1.00
	4 m	-1.122	1.00	1.588	0.99	1.00
$\{\Gamma_{13}, \Gamma_{33}\}$	1 m	9.085	0.00	9.069	0.00	0.0
	2 m	0.475	0.31	0.488	0.54	0.66
	3 m	-0.036	1.00	0.231	0.98	0.99
	4 m	-0.127	1.00	0.221	0.99	0.96
	5 m	-0.186	1.00	0.227	1.00	0.96

The performance of the hypothesis tests is often evaluated using so-called receiver-operating characteristic (ROC) curves (these curves have their origin in the application of radar detection of signals in background noise, but now they are used anywhere such tests are applied, particularly in biomedical decision making; Centor, 1991). The ROC curves relate the power of the test, $p = 1 - \beta$, to the probability of a Type-I error, α. For Setup A, the power of the test is also the probability of detection (the complement of the probability of a miss), and for Setup B, the power is the probability of no detection (the complement of the probability of a false alarm).

Figure 15 shows the ROC curves for Setup A and different anomaly depths, as well as different observation sets (the corresponding figure for Setup B is similar, but not shown). The power of the test increases as the ROC curve approaches the upper left corner of the plot. As expected, the tests are most powerful when the signal is strongest compared to the background noise, and when more than one gradient is observed. The curves are much more sensitive to errors in the statistics of the filter output if the signal-to-noise ratio is small.

Figure 16 compares the tests for the two hypothesis setups, showing that for typical (small) adopted significance levels (e.g., $\alpha = 0.05$), the test under Setup A is more powerful than the test under Setup B. That is, all things being equal, the probability of a miss is smaller than that of a false alarm. This is due to the asymmetry of the ROC curves, which derives from the fact that the shape of the probability density of the maximum filter output depends on the assumed presence or absence of the signal. That is, if the signal is strongly present, the density is Gaussian; if the signal is absent, the density is distinctly asymmetric (non-Gaussian). This result is borne out by the simulations, as well (e.g., compare the last columns of Tables 9 and 10).

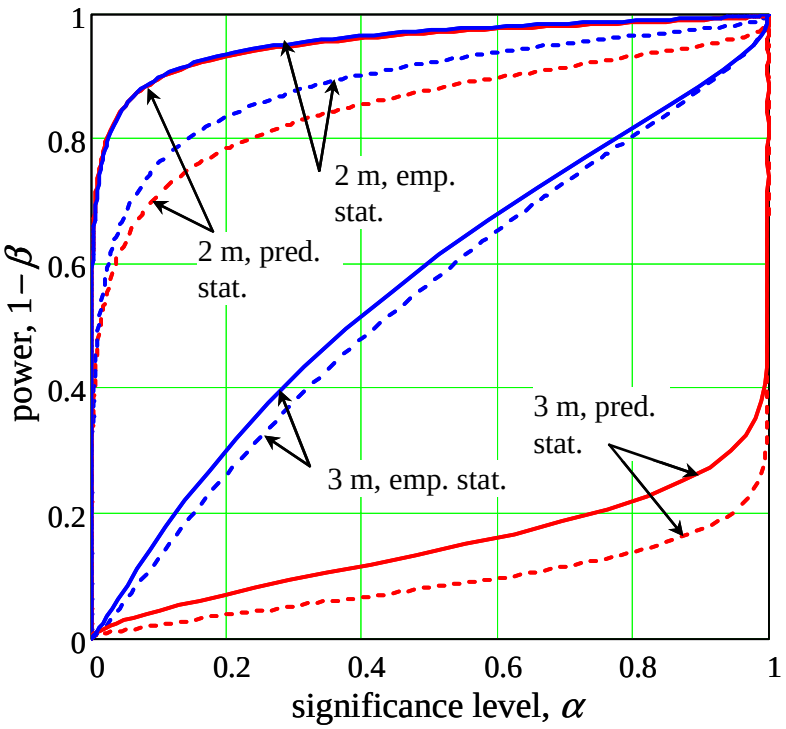


Figure 15: ROC curves under Setup A for anomalies at depths 2 m and 3 m, using predicted and empirical statistics for the filter output. Dashed curves are for the observations $\{\Gamma_{33}\}$; solid curves are for the observations $\{\Gamma_{13}, \Gamma_{33}\}$.

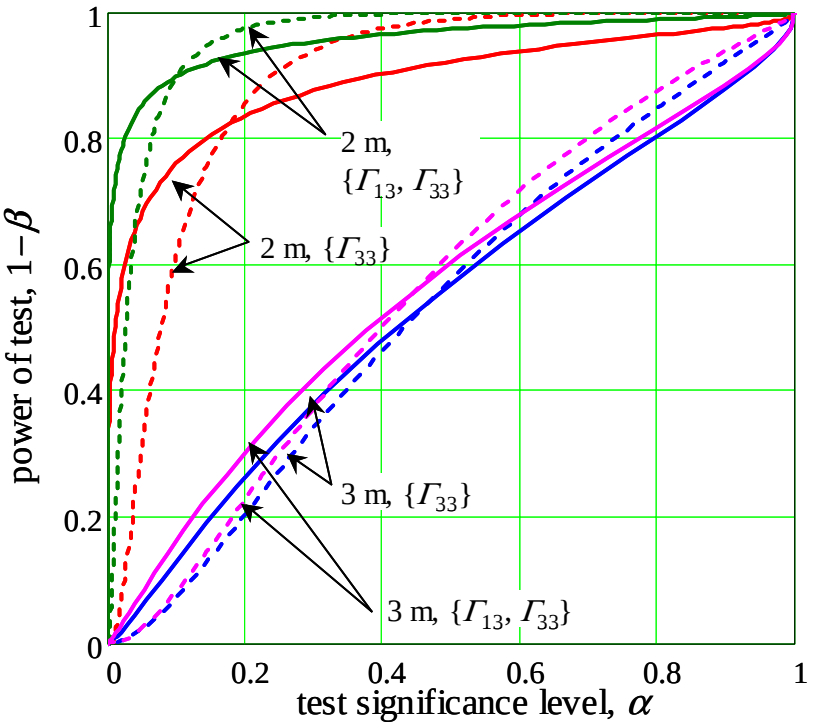


Figure 16: ROC curves for anomalies at depths 2 m and 3 m, using empirical statistics for the filter output, and based on observations $\{\Gamma_{33}\}$ and $\{\Gamma_{13}, \Gamma_{33}\}$. Solid curves correspond to Setup A and dashed curves correspond to Setup B.

XI Summary

The theory was presented for the detection of a linear subsurface anomaly using the matched filter applied to observations of gravity gradients that include correlated background noise and the putative signal of the anomaly. The anomaly is assumed to be a long, right-rectangular prism of negative density contrast with a given (much smaller) cross-section and located at a given depth and orientation. Various cases were considered, including single or multiple observations per point along a single track or multiple parallel tracks. Also, the tracks could be either orthogonal to the anomaly or at an arbitrary azimuth. The maximum output of the matched filter, which presumably locates the signal, is also a test statistic for the Neyman-Pearson hypothesis test of the existence of the signal. Thus, using the statistics for the filter output one can test whether the indicated location is indeed the true location of the anomaly.

Numerical examples of simulated gradients demonstrated the algorithms and the assignment of probabilities of a miss and a false alarm. The examples verified the more general conclusion derived from extensive Monte Carlo simulations (see below) that the calculation of the probability of a miss is more reliable than that of the probability of a false alarm. It was also shown that the filter output is highly correlated in the cross-track direction due to the invariance of the assumed signal in that direction, and that the statistical analysis in this case is not applicable.

Two setups of the Neyman-Pearson hypothesis testing were analyzed. Under Setup A, the null hypothesis posits no anomaly; under Setup B, the null hypothesis assumes that the anomaly

exists (at a given depth and orientation). By simulation analyses, using 1000 realizations of the background noise, it was verified that the test under Setup A is generally more powerful than the test under Setup B, a conclusion that is predicted by corresponding receiver-operating-characteristic (ROC) curves. In general, the decisions to identify the location of the anomaly among the 1000 simulated cases were more robust (accurate) based on the threshold associated with hypothesis Setup A. Still, under Setup B, designed to yield the probability of a false alarm (*POF*), the theoretical and calculated *POF* increased appropriate to the nearly indiscriminate acceptance of the signal-existence hypothesis.

From the simulations as well as the ROC curves, we find, as expected, that correct identification of the signal location improves with added observations and with an increase in the signal-to-noise ratio (shallower anomaly or reduced background noise). We showed that care must be exercised when assigning statistics to the output of the matched filter under the competing hypotheses of the Neyman-Pearson test. Absent a signal in the background field, the *maximum* output of the filter has the statistics of an ordered random variable, which are significantly different from the statistics of the arbitrary output of the filter. This causes some asymmetry in the hypothesis setups and a difference in the power of the corresponding tests.

Finally we observed that multiple parallel tracks are not amenable to the statistical testing since the prerequisites for determining the correct statistics of the order statistic are not fulfilled (independent distributions for the filter output). Moreover, the 2-D matched filter formulation does not perform well if the anomaly is oriented obliquely with respect to multiple parallel tracks. Therefore, it is recommended that for the present application these methods (matched filter and corresponding probability computations) be limited to single tracks of data (but multiple data types per observation point are allowed and beneficial).

References

Butler, D.K. (1984): Microgravimetric and gravity gradient techniques for detection of subsurface cavities, *Geophysics*, 49, No 7, 1084-1096.

Centor. R.M. (1991): Signal detectability: The use of ROC curves and their analyses. *Medical Decision Making*, **11**, 102-106.

Driscoll, M., W.G. Heller, and J.V. White (1990): Multisensor detection of tunnels and other concealed features in a military environment. Tech. Report TR-5859-1, prepared by The Analytic Sciences Corporation for U.S. Army Engineer Waterways Experiment Station, Vicksburg, Mississippi.

Gradshteyn, I.S. and I.M. Ryzhik (1980): *Table of Integrals, Series, and Products*. Academic Press, New York.

David, H.A. (1981): *Order Statistics, Second Edition*. Wiley, Inc., New York.

Dumrongchai, P. (2007): Small anomalous mass detection from airborne gradiometry. Report no.482, Geodetic Science and Surveying, Ohio State University, Columbus, Ohio, http://www.geology.osu.edu/~jekeli.1/OSUReports/.

Forsberg, R. (1987): A New Covariance Model, for Inertial Gravimetry and Gradiometry. *J. Geophys. Res.*, **92**(B2), 1305-1310.

Heller, W.G., S.K. Jordan (1979): Attenuated White Noise Statistical Gravity model. *J. Geophys. Res.*, **84**(B9), 4680-4688.

Jekeli, C. (2003): Statistical analysis of moving-base gravimetry and gravity gradiometry. Report no.466, Geodetic Science, Ohio State University, Columbus, Ohio. http://www.geology.osu.edu/~jekeli.1/OSUReports/reports/report_466.pdf.

Jekeli, C. (2010): Correlation Modeling of the Geopotential Field. In: Handbook of Geomathematics, W. Freeden et al. (eds.), in press.

Meglich, T.M., M.C. Williams, K. Hanna (2005): Subsurface imaging of lava tubes: roadway applications. Report no. FHWA-CFL/TD-05-005, Federal Highway Administration, Lakewood, CO.

Munk, J., R.A. Sheets (1997): Detection of underground voids in Ohio by use of geophysical methods. Report No. FHWA/OH-97/010, prepared by the U.S. Geological Survey under sponsorship from the Ohio Department of Transportation, Columbus, OH.

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, *Journal of Physics D: Applied Physics*, 34, 433-443.

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.

Appendix A – Formulas for the Gravitational Gradient due to Right Rectangular Prism

A.1 Preliminary Considerations

In accord with Newton's law of gravitation, the gravitational potential due to a closed volume, *v*, with density, ρ, is given by

$$V(\boldsymbol{x}) = G\iiint_{v} \frac{\rho(\boldsymbol{x}')}{|\boldsymbol{x}-\boldsymbol{x}'|} dx_1' dx_2' dx_3', \tag{A.1}$$

where *G* is Newton's gravitational constant and $\boldsymbol{x}$ and $\boldsymbol{x}'$ are vectors of (for convenience) Cartesian coordinates; e.g., $\boldsymbol{x} = (x_1, x_2, x_3)^{\mathrm{T}}$. We will assume that the density is a constant: $\rho(\boldsymbol{x}') = \rho_0$, and that the volume *v* is a rectangular block with faces parallel to the coordinate-planes (Figure A.1). With constant limits for the edges of the prism and units chosen so that $G\rho_0 = 1$, the formula (A.1) becomes more explicitly:

$$V(\boldsymbol{x}) = \int_{x_1'=a_1}^{a_2} \int_{x_2'=b_1}^{b_2} \int_{x_3'=c_1}^{c_2} \frac{1}{|\boldsymbol{x}-\boldsymbol{x}'|} dx_1' dx_2' dx_3', \tag{A.2}$$

where the sides of the prism have lengths, $a_2 - a_1$, $b_2 - b_1$, and $c_2 - c_1$, respectively. Because the limits are constants, we may treat the integral as indefinite and apply the limits at the end.

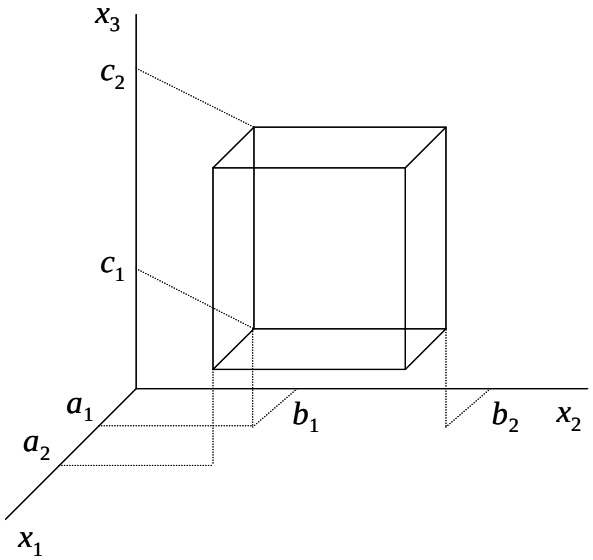


Figure A.1: Geometry of the right-rectangular prism.

For the derivation of the second-order gradients of the potential, *V*, we need one essential result. Noting that

$$\left|\boldsymbol{x}-\boldsymbol{x}'\right| = \sqrt{\left(x_1 - x_1'\right)^2 + \left(x_2 - x_2'\right)^2 + \left(x_3 - x_3'\right)^2}\,, \tag{A.3}$$

we find

$$\frac{\partial}{\partial x_j}\frac{1}{\left|\boldsymbol{x}-\boldsymbol{x}'\right|} = -\frac{\partial}{\partial x_j'}\frac{1}{\left|\boldsymbol{x}-\boldsymbol{x}'\right|}, \quad j = 1,2,3\,. \tag{A.4}$$

With this we are ready to derive the gravitational gradients:

$$\Gamma_{jk} = \frac{\partial^2 V}{\partial x_j \partial x_k}, \quad j,k = 1,2,3\,. \tag{A.5}$$

Because of the symmetry of the kernel function, $1/\left|\boldsymbol{x}-\boldsymbol{x}'\right|$, and the constant integration limits, we need to consider only two gradients, a cross-gradient, Γ_{12}, and an in-line gradient, Γ_{11}. Once formulas have been derived for these gradients, the other cross-gradients and in-line gradients can be obtained by cyclically permuting the indices on the coordinates. Since the gradients are not continuous on the mass boundary, we will always assume that $\boldsymbol{x} \neq \boldsymbol{x}'$; that is, the computation point is never on (or inside) the volume surface.

A.2 Reduction to Single and Double Integrals

From equations (A.2) and (A.4), we have:

$$\frac{\partial V}{\partial x_1} = \int_{x_2'=b_1}^{b_2}\int_{x_3'=c_1}^{c_2} -\frac{1}{\left|\boldsymbol{x}-\boldsymbol{x}'\right|}\Bigg|_{x_1'=a_1}^{a_2} dx_3'\,dx_2'\,. \tag{A.6}$$

Then, again using equation (A.4), the cross-gradient, Γ_{12}, is

$$\frac{\partial^2 V}{\partial x_1 \partial x_2} = \int_{x_3'=c_1}^{c_2} \frac{1}{\left|\boldsymbol{x}-\boldsymbol{x}'\right|}\Bigg|_{x_1'=a_1}^{a_2}\Bigg|_{x_2'=b_1}^{b_2} dx_3'\,. \tag{A.7}$$

The in-line gradient, Γ_{11}, is

$$\frac{\partial^2 V}{\partial x_1^2} = \int_{x_2'=b_1}^{b_2}\int_{x_3'=c_1}^{c_2} \frac{\left(x_1 - x_1'\right)}{\left|\boldsymbol{x}-\boldsymbol{x}'\right|^3}\Bigg|_{x_1'=a_1}^{a_2} dx_3'\,dx_2'\,. \tag{A.8}$$

Simple analytic expressions are available for integrals (A.7) and (A.8), as seen in the next Section.

A.3 Anti-Derivatives

Because the integration limits are constants, these integrals may be considered as indefinite until the end. Furthermore, to simplify the evaluation we may temporarily define dummy variables

$$x = x_1 - x_1', \quad y = x_2 - x_2', \quad z = x_3 - x_3'. \tag{A.9}$$

Then, the integrals (A.7) and (A.8) are of the following two types:

$$I = \int_z \frac{1}{\sqrt{x^2 + y^2 + z^2}}\, dz, \qquad J = \int_y \int_z \frac{1}{\left(x^2 + y^2 + z^2\right)^{3/2}}\, dz dy, \tag{A.10}$$

The analytic evaluation of J with respect to z is given by Gradshteyn and Ryzhik (1980, 2.271-5):

$$\int_z \frac{1}{\left(x^2 + y^2 + z^2\right)^{3/2}}\, dz = \frac{z}{\left(x^2 + y^2\right)\sqrt{x^2 + y^2 + z^2}}. \tag{A.11}$$

Now, factoring the first quadratic in the denominator of equation (A.11), we find

$$\frac{1}{x^2 + y^2} = \frac{\frac{1}{2x}}{x - iy} + \frac{\frac{1}{2x}}{x + iy}. \tag{A.12}$$

Hence,

$$J = \frac{z}{2x}\int_y \frac{1}{(x - iy)\sqrt{x^2 + y^2 + z^2}}\, dy + \frac{z}{2x}\int_y \frac{1}{(x + iy)\sqrt{x^2 + y^2 + z^2}}\, dy. \tag{A.13}$$

Changing variables of integration to $u = x \pm iy$, equation (A.13) becomes

$$J = -\frac{z}{i2x}\int_y \frac{1}{u\sqrt{x^2 - (u - x)^2 + z^2}}\, du + \frac{z}{i2x}\int_y \frac{1}{u\sqrt{x^2 - (u - x)^2 + z^2}}\, du. \tag{A.14}$$

The integrals do not cancel because the limits of integration are different.

We now define

$$\begin{aligned} I_{-1}(\alpha, \beta, \gamma) &= \int \frac{1}{x\sqrt{\alpha + 2\beta x + \gamma x^2}}\, dx \\ I_0(\alpha, \beta, \gamma) &= \int \frac{1}{\sqrt{\alpha + 2\beta x + \gamma x^2}}\, dx \end{aligned} \tag{A.15}$$

From Gradshteyn and Ryzhik (1980, 2.266), we have (leaving out constants)

$$I_{-1}(\alpha,\beta,\gamma)=\begin{cases}-\dfrac{1}{\sqrt{\alpha}}\ln\dfrac{\alpha+\beta x+\sqrt{\alpha\left(\alpha+2\beta x+\gamma x^2\right)}}{x}, & \alpha\neq 0,\ \ \beta^2-\alpha\gamma\neq 0;\\ -\dfrac{1}{\beta x}\sqrt{2\beta x+\gamma x^2}, & \alpha=0,\qquad \beta\neq 0;\\ \dfrac{1}{\sqrt{\alpha}}\ln\dfrac{x}{\alpha+\beta x} & \alpha\neq 0,\ \ \beta^2-\alpha\gamma=0.\end{cases} \tag{A.16}$$

And, from Gradshteyn and Ryzhik (1980, 2.261), we have (again, leaving out constants)

$$I_0(\alpha,\beta,\gamma)=\begin{cases}\dfrac{1}{\sqrt{\gamma}}\ln\left(\beta+\gamma x+\sqrt{\gamma\left(\alpha+2\beta x+\gamma x^2\right)}\right), & \gamma\neq 0,\ \ \beta^2-\alpha\gamma\neq 0;\\ \dfrac{1}{\beta}\sqrt{\alpha+2\beta x}, & \gamma=0,\qquad \beta\neq 0;\\ \dfrac{1}{\sqrt{\gamma}}\ln\left(\beta+\gamma x\right), & \gamma\neq 0,\ \ \beta^2-\alpha\gamma=0.\end{cases} \tag{A.17}$$

Hence, the integral in *J* becomes

$$\int\frac{1}{u\sqrt{x^2-(u-x)^2+z^2}}\,du=-\frac{1}{z}\ln\frac{z^2+xu+z\sqrt{x^2-(u-x)^2+z^2}}{u}. \tag{A.18}$$

Now substituting *y* back into these solutions, we find

$$J=\frac{1}{i2x}\ln\frac{z^2+x^2-ixy+z\sqrt{x^2+y^2+z^2}}{x-iy}-\frac{1}{i2x}\ln\frac{z^2+x^2+ixy+z\sqrt{x^2+y^2+z^2}}{x+iy}. \tag{A.19}$$

With some rearrangement of terms we also have

$$J=\frac{1}{i2x}\ln\frac{x+iy}{x-iy}-\frac{1}{i2x}\ln\frac{z^2+x^2+z\sqrt{x^2+y^2+z^2}+ixy}{z^2+x^2+z\sqrt{x^2+y^2+z^2}-ixy}. \tag{A.20}$$

Using

$$\tan^{-1}w=\frac{1}{2i}\ln\frac{1+iw}{1-iw}, \tag{A.21}$$

we obtain

$$J = \frac{1}{x}\tan^{-1}\frac{y}{x} - \frac{1}{x}\tan^{-1}\frac{xy}{z^2 + x^2 + zr}, \tag{A.22}$$

where $r = \sqrt{x^2 + y^2 + z^2}$. These inverse tangents can be combined using the addition formula for tangents, yielding

$$J = \frac{1}{x}\tan^{-1}\frac{yz}{xr}. \tag{A.23}$$

Noting that $dzdy = dx_3'dx_2'$, we see that by combining equations (A.8), (A.9), (A.10), and (A.23), we finally obtain

$$\Gamma_{11} = \tan^{-1}\frac{(x_2 - x_2')(x_3 - x_3')}{(x_1 - x_1')r}\Bigg|_{x_1'=a_1}^{a_2}\Bigg|_{x_2'=b_1}^{b_2}\Bigg|_{x_3'=c_1}^{c_2}, \tag{A.24}$$

where now

$$r = \sqrt{(x_1 - x_1')^2 + (x_2 - x_2')^2 + (x_3 - x_3')^2}. \tag{A.25}$$

Applying equation (A.17) to *I*, defined in equation (A.12), we have

$$I = \ln\left(z + \sqrt{x^2 + y^2 + z^2}\right). \tag{A.26}$$

Again, using (9) and noting that $dz = -dx_3'$, the cross-gradient, equation (A.7), is seen to be

$$\Gamma_{12} = -\ln(x_3 - x_3' + r)\Big|_{x_1'=a_1}^{a_2}\Bigg|_{x_2'=b_1}^{b_2}\Bigg|_{x_3'=c_1}^{c_2}, \tag{A.27}$$

where *r* is given by equation (A.25).

Clearly, by the symmetry already mentioned and the constant integration limits, the remaining gradients follow by cyclic permutation of the coordinates. In summary, we have

$$\Gamma_{11} = \tan^{-1}\frac{(x_2 - x_2')(x_3 - x_3')}{(x_1 - x_1')r}\Bigg|_{x_1'=a_1}^{a_2}\Bigg|_{x_2'=b_1}^{b_2}\Bigg|_{x_3'=c_1}^{c_2}, \tag{A.28}$$

$$\Gamma_{22} = \tan^{-1} \frac{(x_3 - x_3')(x_1 - x_1')}{(x_2 - x_2')r} \Bigg|_{x_1'=a_1}^{a_2} \Bigg|_{x_2'=b_1}^{b_2} \Bigg|_{x_3'=c_1}^{c_2}, \tag{A.29}$$

$$\Gamma_{33} = \tan^{-1} \frac{(x_1 - x_1')(x_2 - x_2')}{(x_3 - x_3')r} \Bigg|_{x_1'=a_1}^{a_2} \Bigg|_{x_2'=b_1}^{b_2} \Bigg|_{x_3'=c_1}^{c_2}, \tag{A.30}$$

$$\Gamma_{12} = -\ln(x_3 - x_3' + r) \Big|_{x_1'=a_1}^{a_2} \Big|_{x_2'=b_1}^{b_2} \Big|_{x_3'=c_1}^{c_2}, \tag{A.31}$$

$$\Gamma_{23} = -\ln(x_1 - x_1' + r) \Big|_{x_1'=a_1}^{a_2} \Big|_{x_2'=b_1}^{b_2} \Big|_{x_3'=c_1}^{c_2}, \tag{A.32}$$

$$\Gamma_{31} = -\ln(x_2 - x_2' + r) \Big|_{x_1'=a_1}^{a_2} \Big|_{x_2'=b_1}^{b_2} \Big|_{x_3'=c_1}^{c_2}. \tag{A.33}$$

And, of course, $\Gamma_{13} = \Gamma_{31}$, $\Gamma_{21} = \Gamma_{12}$, and $\Gamma_{32} = \Gamma_{23}$.

Appendix B – Gradients for an Arbitrarily Oriented Prism

The previous Appendix derived simple formulas for the gravitational gradients due to a rectangular prism, provided the edges of the prism are parallel to the coordinate axes. But suppose the prism is rotated arbitrarily with respect to these axes.

Let $\boldsymbol{x} = (x_1, x_2, x_3)$ be the system of coordinates in which we wish to compute the gravitational gradients due to the prism. We assume for the moment that the prism has its center at the coordinate origin. Let $\boldsymbol{u} = (u_1, u_2, u_3)$ be the system obtained by rotating the $\boldsymbol{x}$-system so that the axes are now parallel to the prism. Let C_x^u be the rotation matrix that describes this rotation.

We can compute the gravitational gradients, Γ^u, at an point in the $\boldsymbol{u}$-system using equations (A.28) through (A.33), replacing $\boldsymbol{x} = (x_1, x_2, x_3)$ by $\boldsymbol{u} = (u_1, u_2, u_3)$ and assuming that the limits of integration refer to the $\boldsymbol{u}$-system. Then, the gradients for the corresponding point in the $\boldsymbol{x}$-system are determined by the following standard transformation of a tensor from one system ($\boldsymbol{u}$-system) to another ($\boldsymbol{x}$-system):

$$\Gamma^x = C_u^x \Gamma^u C_x^u . \tag{B.1}$$

This transformation refers to changes in directions of derivatives at a point.

More specifically, we have a right, rectangular prism oriented by three angles with respect to the $\boldsymbol{x}$-system. A dip angle, ξ, from the horizontal, an azimuth, α, from north, and a rotation, η, about the longitudinal axis of the prism, as seen in Figure B.1. That is, we start with the prism with its long axis parallel to the x_1-axis and centered on the (x_2, x_3)-plane. This is rotated by the azimuth angle, α; then dipped by the angle, ξ; and finally twisted by the angle, η. The coordinate system, (u_1, u_2, u_3), defined by the axes of the prism is thus obtained by the rotation (following the rotations as defined in Figure 1):

$$\boldsymbol{u} = R_1(-\eta) R_2(-\xi) R_3(-\alpha) \boldsymbol{x} = C_x^u \boldsymbol{x} . \tag{B.2}$$

In order to compute the gradients of the rotated prism in the $\boldsymbol{x}$ system, we first compute the coordinates, (u_1, u_2, u_3), given the coordinates (x_1, x_2, x_3), according to the transformation (B.2). Then, we compute the gradients in the $\boldsymbol{u}$-system at that point (using parameters for the prism defined in the $\boldsymbol{u}$-system). Finally, we change directions of the derivatives at the computation point by applying the transformation (B.1).

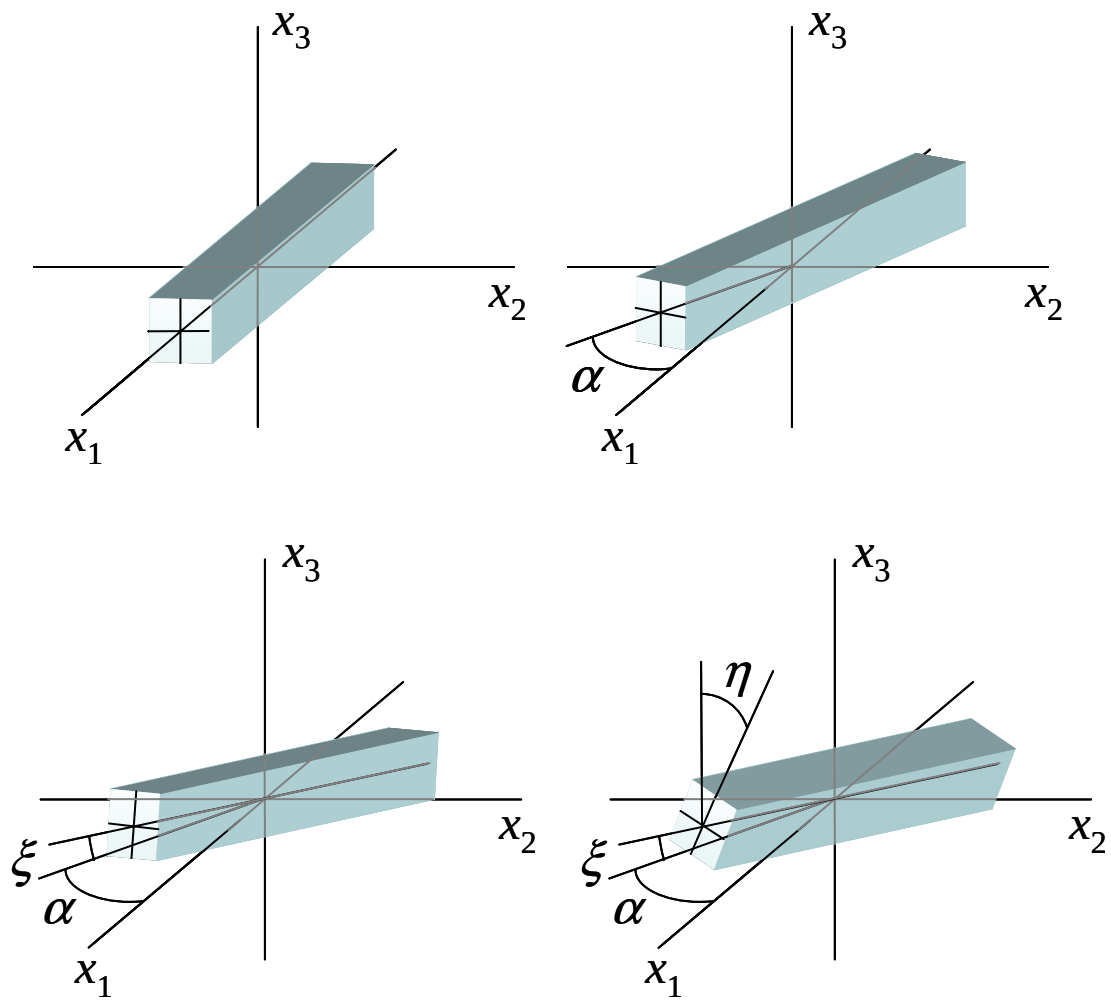


Figure B.1: Original prism aligned with coordinate axes and subsequent rotation by azimuth, α, dip, ξ, and twist, η.

For a prism not centered at the origin of the $\boldsymbol{x}$-system, we apply the alternative transformation to the given $\boldsymbol{x}$-system coordinates:

$$\boldsymbol{u} = C_x^u \left(\boldsymbol{x} - \boldsymbol{x}_0 \right), \tag{B.3}$$

where $\boldsymbol{x}_0$ is the location of the center of the prism. For these coordinates, $\left(u_1, u_2, u_3\right)$, we compute the gradients (again, using limits of integration in the $\boldsymbol{u}$-system, where the prism is centered at the origin of the $\boldsymbol{u}$-system). Then we apply the transformation (B.1) to obtain the gradients in the $\boldsymbol{x}$-system. Note that transformation (B.3) is used to obtain the needed coordinates of the computation point in the $\boldsymbol{u}$-system, while transformation (B.1) suffices to transform the directions of the derivatives, irrespective of $\boldsymbol{x}_0$.

Appendix C – The Test Statistic for the Matched Filter

With Gaussian noise, the observations (*N* in number) have the following joint density:

$$f_{\mathbf{z}}(\mathbf{z}) = \frac{1}{\sqrt{|\phi|}(2\pi)^{N/2}} e^{-\frac{1}{2}\sum_j\sum_k (\mathbf{z}_j - \mathbf{s}_{j-m})^{\mathrm{T}} \phi^{-1}(j,k)(\mathbf{z}_k - \mathbf{s}_{k-m})} = \frac{1}{\sqrt{|\phi|}(2\pi)^{N/2}} e^{-\frac{1}{2}(\mathbf{z}-\mathbf{s}_{-m})^{\mathrm{T}} \phi^{-1}(\mathbf{z}-\mathbf{s}_{-m})}. \quad \text{(C.1)}$$

where $\mathbf{s}_{-m}$ is the vector of signals displaced by *m*. For the hypotheses of Setup A (hypotheses (57) and (58)), the likelihood ratio is

$$LR(x_m) = \frac{f_{\mathbf{z}|H_1}(\mathbf{z} \mid H_1)}{f_{\mathbf{z}|H_0}(\mathbf{z} \mid H_0)} = e^{-\frac{1}{2}(\mathbf{z}-\bar{\mathbf{s}}_{-m})^{\mathrm{T}} \phi^{-1}(\mathbf{z}-\bar{\mathbf{s}}_{-m}) + \frac{1}{2}\mathbf{z}^{\mathrm{T}}\phi^{-1}\mathbf{z}}. \quad \text{(C.2)}$$

To simplify, consider the natural logarithm of this expression, since the right side is a monotonic positive function:

$$\ln(LR_m) = \bar{\mathbf{s}}_{-m}^{\mathrm{T}} \phi^{-1} \mathbf{z} - \frac{1}{2} \bar{\mathbf{s}}_{-m}^{\mathrm{T}} \phi^{-1} \bar{\mathbf{s}}_{-m}. \quad \text{(C.3)}$$

By reversing the signs on the coordinates, shifting the origin by *m*, and noting that the noise is stationary, it is readily shown with equations (30) and (29) that

$$\ln(LR_m) = \lambda \bar{y} - \frac{1}{2}\lambda^2, \quad \text{(C.4)}$$

where $\bar{y} \equiv y_m$. Clearly, the same result holds for Setup B (hypotheses (59) and (60)), provided the *LR* is reversed.

Appendix D – Covariance Matrix of the Matched Filter Output

The covariance of the output is defined as (we consider only a single track of data with one type of observation at each data point)

$$\operatorname{cov}(y_r, y_{r'}) = \mathcal{E}\left((y_r - \mu_r)(y_{r'} - \mu_{r'})\right) \tag{D.1}$$

where the mean, μ_r, is given by equation (61). Combining equations (30) and (61), we simplify, in view of equation (3),

$$y_r - \mu_r = \frac{1}{\lambda}\sum_j \sum_k \overline{s}_{-k}\phi_-^{-1}(j,k)\, n_{r-j}\,. \tag{D.2}$$

Then

$$\begin{aligned}\operatorname{cov}(y_r, y_{r'}) &= \mathcal{E}\left(\frac{1}{\lambda^2}\sum_j\sum_k\sum_{j'}\sum_{k'}\overline{s}_{-k}\phi_-^{-1}(j,k)\, n_{r-j}\overline{s}_{-k'}\phi_-^{-1}(j',k')\, n_{r'-j'}\right)\\ &= \frac{1}{\lambda^2}\sum_j\sum_k\sum_{j'}\sum_{k'}\overline{s}_{-k}\phi_-^{-1}(j,k)\,\phi_-(r'-j', r-j)\,\phi_-^{-1}(j',k')\,\overline{s}_{-k'}\end{aligned} \tag{D.3}$$

which is independent of the actual signal in the observations.

If the noise is uncorrelated, e.g., $\phi = I$ (the identity matrix), then

$$\begin{aligned}\operatorname{cov}(y_r, y_{r'}) &= \frac{1}{\lambda^2}\sum_j\sum_k\sum_{j'}\overline{s}_{-k}\phi_-^{-1}(j,k)\,\phi_-(r'-j', r-j)\sum_{k'}\phi_-^{-1}(j',k')\,\overline{s}_{-k'}\\ &= \frac{1}{\lambda^2}\sum_j\sum_k\overline{s}_{-k}\phi_-^{-1}(j,k)\sum_{j'}\phi_-(r'-j', r-j)\,\overline{s}_{-j'}\\ &= \frac{1}{\lambda^2}\sum_j\overline{s}_{r-r'-j}\sum_k\overline{s}_{-k}\phi_-^{-1}(j,k)\\ &= \frac{1}{\lambda^2}\sum_j\overline{s}_{r-r'-j}\overline{s}_{-j}\end{aligned} \tag{D.4}$$

which shows that the output is uncorrelated in this case only if the assumed signal is also very localized. Specifically, if it is the delta (impulse) function,

$$\overline{s}_j = \delta_j = \begin{cases}1, & j = 0\\ 0, & j \neq 0\end{cases} \tag{D.5}$$

then, since $\lambda^2 = 1$,

$$\operatorname{cov}\left(y_r, y_{r'}\right) = \sum_j \delta_{r-r'-j}\delta_{-j} = \delta_{r-r'} = \begin{cases} 1, & r = r' \\ 0, & r \neq r' \end{cases} \tag{D.6}$$

The variance in the general case is ($r = r'$ in equation (D.3)):

$$\begin{aligned} \operatorname{var}\left(y_r\right) &= \frac{1}{\lambda^2}\sum_j\sum_k\sum_{j'}\sum_{k'} \overline{s}_{-k}\phi_-^{-1}\left(j,k\right)\phi_-\left(-j',-j\right)\phi_-^{-1}\left(j',k'\right)\overline{s}_{-k'} \\ &= \frac{1}{\lambda^2}\sum_j\sum_k\sum_{j'}\sum_{k'} \overline{s}_{-k}\phi_-^{-1}\left(j,k\right)\phi_-\left(j,j'\right)\phi_-^{-1}\left(j',k'\right)\overline{s}_{-k'} \\ &= \frac{1}{\lambda^2}\sum_j\sum_k\sum_{k'} \overline{s}_{-k}\phi_-^{-1}\left(j,k\right)\overline{s}_{-k'}\sum_{j'}\phi_-\left(j,j'\right)\phi_-^{-1}\left(j',k'\right) \end{aligned} \tag{D.7}$$

where

$$\sum_{j'}\phi_-\left(j,j'\right)\phi_-^{-1}\left(j',k'\right) = \begin{cases} 0, & j \neq k' \\ 1, & j = k' \end{cases} \tag{D.8}$$

Hence

$$\begin{aligned} \operatorname{var}\left(y_r\right) &= \frac{1}{\lambda^2}\sum_j\sum_k \overline{s}_{-k}\phi_-^{-1}\left(j,k\right)\overline{s}_{-j} \\ &= 1 \end{aligned} \tag{D.9}$$

in view of equation (26). We see that the variance is independent of *r*, the location of the filter output.

For the signal modeled as an impulse (equation (D.5)), the covariance of the output, equation (D.3) becomes

$$\begin{aligned} \operatorname{cov}\left(y_r, y_{r'}\right) &= \frac{1}{\lambda^2}\sum_j\sum_k\sum_{j'}\sum_{k'} \delta_{-k}\phi_-^{-1}\left(j,k\right)\phi_-\left(r'-j',r-j\right)\phi_-^{-1}\left(j',k'\right)\delta_{-k'} \\ &= \frac{1}{\lambda^2}\sum_j\sum_{j'} \phi_-^{-1}\left(j,0\right)\phi_-\left(j-r, j'-r'\right)\phi_-^{-1}\left(j',0\right) \end{aligned} \tag{D.10}$$

Again, we have with equation (D.8) and $\lambda^2 = \phi_-^{-1}\left(0,0\right)$

$$\begin{aligned}\operatorname{cov}\left(y_r, y_{r'}\right) &= \frac{1}{\lambda^2}\sum_j \phi_-^{-1}\left(j,0\right)\sum_{j'}\phi_-\left(j-r+r', j'\right)\phi_-^{-1}\left(j',0\right) \\ &= \frac{1}{\lambda^2}\sum_j \phi_-^{-1}\left(j,0\right)\begin{cases}0, & j-r+r'\neq 0\\ 1, & j-r+r'=0\end{cases} \\ &= \frac{\phi_-^{-1}\left(r-r',0\right)}{\phi_-^{-1}\left(0,0\right)}\end{aligned} \tag{D.11}$$

We see again that the output is correlated unless both ϕ is proportional to the identity function and the signal is the delta function.

Appendix E – Gradient Field Realizations

The background gravity field was realized from a correlated two-dimensional random process on the plane. The mutual consistency of different second-order gradients must be embedded in the construction of the covariance matrix, ϕ, appearing in equations (28) and (30), and can be ensured by designing an appropriate covariance model or, equivalently, a power spectral density (psd) model for the *potential, W,* (more precisely, a residual potential excluding long-wavelength features) of the gravity field. We used the model described by Jekeli (2003), who also showed how to construct corresponding, mutually consistent models (for both psd's and covariance functions) for the gravity gradients. Similar models were developed, e.g., by Heller and Jordan (1979) and Forsberg (1987). The covariance matrix thus constructed is invertible, as required in equation (27), since the model is positive definite by design (the psd is positive for all frequencies).

It is noted that in the case of observations with correlated noise, the matched filter often is applied only after the observations are whitened. Such a procedure was employed in the case of gravity gradients by Driscoll et al. (1990). This simplifies the filter equations, but in essence requires a determination of the covariance of the field. Since our background gravitational field is simulated from a given covariance model, we need not perform this two-step process and can apply the matched filter directly to the correlated observations according to equation (30). In this case, ϕ is the sum of the covariance model for the correlated background gravitational field and the variance of the (uncorrelated) measurement noise (equation (4)).

If *T* is the (disturbing) gravitational potential, *T*, a particular (disturbing) gradient in a local Cartesian coordinate system is given by

$$\Gamma_{jk} = \frac{\partial^2 T}{\partial x_j \partial x_k} . \tag{E.1}$$

Therefore, given a model for the psd of *T*, the psd of a gradient, $\Phi_\Gamma(n_1, n_2)$, is readily modeled using the transform of the corresponding derivative operator in the frequency domain (Jekeli, 2003).

We wish to generate an $M_1 \times M_2$ grid of gradients ($M_1 = M_2 = 100$) with data spacing, Δx_1 and Δx_2 ($\Delta x_1 = \Delta x_2 = 1$ m). The *spectrum* of the gradient thus depends on the 2-D frequency vector, $\left(n_1/(M_1\Delta x_1), n_2/(M_2\Delta x_2)\right)$ and may be synthesized on this grid according to

$$G_{n_1,n_2} = \left(b_{n_1,n_2} + ic_{n_1,n_2}\right)\sqrt{M_1\Delta x_1 M_2 \Delta x_2}\sqrt{\Phi_\Gamma(n_1,n_2)}, \tag{E.2}$$

where b_{n_1,n_2} and c_{n_1,n_2} are normally distributed random variables:

$$b_{n_1,n_2} \sim \mathcal{N}\left(0, 1/\sqrt{2}\right), \quad c_{n_1,n_2} \sim \mathcal{N}\left(0, 1/\sqrt{2}\right), \tag{E.3}$$

The standard deviations, $1/\sqrt{2}$, ensure that the standard deviation of the square-root of the periodogram of G_{n_1,n_2} is $\sqrt{\Phi_\Gamma(n_1,n_2)}$, for all n_1, n_2. The inverse discrete Fourier transform of G_{n_1,n_2} thus yields a realization of the corresponding gradient field. For mutual consistency the same random numbers, b_{n_1,n_2} and c_{n_1,n_2}, were used for all gradients belonging to a particular realization of the field.

Figure E.1 shows a synthesis (one realization) of the background gradient, Γ_{33}, on the 100 m $\times$ 100 m grid (1 m spacing). The covariance model used for the gradient has a variance of $(94.2\ \text{E})^2$ (1 E = 1 Eötvös = 10^{-9} s^{-2}), and a correlation length of about 9 km ($1/e$ point). A total of 1000 such realizations were generated to determine the probabilities of success and failure in detecting a particular anomaly buried in this background. Similar results were achieved even with as few as 100 realizations; therefore, we believe that 1000 realizations provide an adequate assessment of the methods.

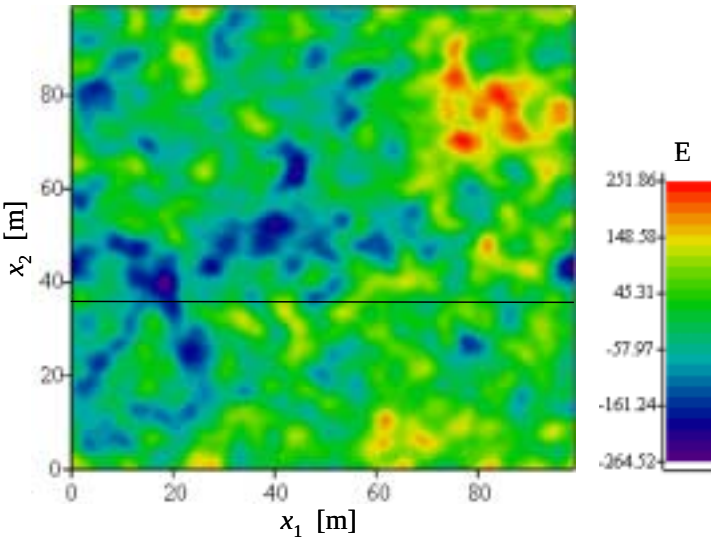


Figure E.1: Vertical-vertical gradients, Γ_{33}, synthesized using a psd model and normally distributed spectral components (equations (E.2) and (E.3)). The horizontal black line indicates the profile used in Figure 14.