Optimizing Techniques and Cramér-Rao Bound for Passive Source Location Estimation

A Thesis Submitted
to the College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
in the Department of Electrical and Computer Engineering
University of Saskatchewan

by

Sha Li

Saskatoon, Saskatchewan, Canada
July 2017

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, it is agreed that the Libraries of this University may make it freely available for inspection. Permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professors who supervised this thesis work or, in their absence, by the Head of the Department of Electrical and Computer Engineering or the Dean of the College of Graduate and Postdoctoral Studies at the University of Saskatchewan. Any copying, publication, or use of this thesis, or parts thereof, for financial gain without the written permission of the author is strictly prohibited. Proper recognition shall be given to the author and to the University of Saskatchewan in any scholarly use which may be made of any material in this thesis.

Request for permission to copy or to make any other use of material in this thesis in whole or in part should be addressed to:

Head of the Department of Electrical and Computer Engineering
57 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan, Canada
S7N 5A9

Dean of the College of Graduate and Postdoctoral Studies
Room 116 Thorvaldson Building
110 Science Place
Saskatoon, Saskatchewan, Canada
S7N 5C9

Abstract

This work is motivated by the problem of locating potential unstable areas in underground potash mines with better accuracy more consistently while introducing minimum extra computational load. It is important for both efficient mine design and safe mining activities, since these unstable areas may experience local, low-intensity earthquakes in the vicinity of an underground mine. The object of this thesis is to present localization algorithms that can deliver the most consistent and accurate estimation results for the application of interest.

As the first step towards the goal, three most representative source localization algorithms given in the literature are studied and compared. A one-step energy-based grid search (EGS) algorithm is selected to address the needs of the application of interest.

The next step is the development of closed-form Cramér-Rao bound (CRB) expressions. The mathematical derivation presented in this work deals with continuous signals using the Karhunen-Loève (K-L) expansion, which makes the derivation applicable to non-stationary Gaussian noise problems. Explicit closed-form CRB expressions are presented only for stationary Gaussian noise cases using the spectrum representation of the signal and noise though.

Using the CRB comparisons, two approaches are proposed to further improve the EGS algorithm. The first approach utilizes the corresponding analytic error estimation variance (EEV) expression presented in the work of Salt and Daku to derive an amplitude weight expression, optimal in terms of minimizing this EEV, for the case of additive Gaussian noise with a common spectrum interpretation across all the sensors. An alternate non-iterative amplitude weighting scheme is proposed based on the optimal amplitude weight expression. It achieves the same performance with less calculation compared with the traditional iterative approach.

The second approach tries to optimize the EGS algorithm in the frequency domain. An analytic frequency weighted EEV expression is derived using spectrum representation and the stochastic process theory. Based on this EEV expression, an integral equation is established and solved using the calculus of variations technique. The solution corresponds to a filter transfer function that is optimal in the sense that it minimizes this analytic frequency domain EEV. When various parts of the frequency domain EEV expression are ignored during the minimization procedure using Cauchy-Schwarz inequality, several different filter transfer functions result. All of them turn out to be well known classical filters that have been developed in the literature and used to deal with source localization problems. This demonstrates that in terms of minimizing the analytic EEV, they are all suboptimal, not optimal.

Monte Carlo simulation is performed and shows that both amplitude and frequency weighting bring obvious improvement over the unweighted EGS estimator.

Acknowledgments

The author wishes to thank God for his amazing grace, power, and presence that enabled her to finalize this research.

The author wishes to thank Prof. Brian Daku for his advice during the course of this work and the preparation of this thesis. The financial assistance was provided by the University of Saskatchewan as a scholarship and Prof. Daku as an assistantship.

The author would also like to thank her parents, Li Xianlin and Zhang Xiurong, for their everlasting support and earnest love. And last but not the least, her sweetheart Vincent Hu, for his understanding, cooperation, and sincere prayers that helped make this thesis possible.

Dedication

To my dearest God,
Mother, Zhang Xiurong,
Father, Li Xianlin,
and Son, Vincent Hu

Table of Contents

List of Figures

List of Abbreviations

AEDA	Adaptive Eigenvalue Decomposition Algorithm
AML	Approxomate Maximum Likelihood
AWEGS	Amplitude Weighting Energy-based Grid Search
CGN	Coloured Gaussian Noise
CRB	Cramér-Rao Bound
DFT	Discrete Fourier Transform
DOA	Direction of Arrival
EEV	Error Estimation Variance
EGS	Energy-based Grid Search
FT	Fourier Transform
FWEGS	Frequency Weighting Energy-based Grid Search
GCC	Generalized Cross Correlation
GPS	Global Positioning System
LMS	Least Mean Square
ML	Maximum Likelihood
PHAT	Phase Transform
RFID	Radio Frequency Identification
RMSE	Root Mean Squared Error
SNR	Signal-to-Noise Ratio
TDOA	Time Difference of Arrival
WGN	White Gaussian Noise

1 Introduction

This chapter starts with some background information on various acoustic source localization applications. Then the specific application and problems under investigation are described. This is followed by the outline of the thesis.

1.1 Background

Acoustic source localization has been investigated in a variety of research areas. It was probably the concern for submarine warfare in the deep ocean that initiated the earliest research in the area of acoustic source localization almost a century ago [1]. Interest has expanded from this deep water application to include coastal water applications that involve a wide range of targets to locate, such as oil spills and other pollution that may cause habitat degradation.

Passive underwater acoustic source localization techniques have also been applied in the study of animal and environmental science to locate whales, dolphins, and other sea creatures that make acoustic sound. It is also applied to locate and monitor undersea volcanic eruptions, earthquakes, and continental collisions in the areas of geophysics, geoscience, and other earth sciences.

In most of the above applications, signals that emanate directly from the source or from the disturbance caused by the source event are usually detected and/or measured using a sensor array deployed underwater in a uniform arrangement. Using the signals measured, ranging of the source/event can be achieved [2–5].

It was within the last few decades that the technology of robotics witnessed a rapid growth in the area of undersea exploration [6]. A well-established technology routinely used in the offshore industry today is the adoption of remotely operated vehicles that are linked to the mother ship by a tether cable. A more recent technology involves autonomous underwater vehicles (AUV) capable of navigating on their own without constant intervention by a human operator. AUVs utilize a Global Positioning System (GPS) and a group of freely floating surface buoys that work together to help the underwater vehicle determine its own absolute coordinates. These AUVs help make continuous long time exploration missions much easier without imposing great distress on people.

Other passive acoustic localization applications have emerged during the last several decades for use on land as well. Locating and tracking a speaker through the use of an array of microphones has received much attention due to the increasing demand for video communication/conferencing over the Internet [7–9]. The goal is to provide the audience, at the other side of the network, with a continuous view of the speaker even when he moves around freely in the venue. Voice quality can be optimized as well, since the microphones will always steer towards the speaker.

A microphone array can be also utilized to locate a mobile robot that works and moves around within a designated area. Another important role a microphone array can have is in the health care industry, for example, remotely monitoring the elderly or in-home patients who might suddenly fall or faint. Source localization and tracking techniques are also useful in navigation systems that help in tracking patient movements in medical wards where knowledge of patient location and movement is critical in providing faster assistance or preemptive action, and prevent patient mix-ups [10].

Indoor and outdoor wireless device and network-centric positioning is another area where source localization techniques are in high demand. The localization accuracy required varies, depending on the application and the technology adopted for it. Most basic location-based mobile services do not require positioning of high accuracy. For example, people are usually satisfied with the weather broadcast for their current city, not their specific neighbourhood. Though, police and medical services demand much more accuracy in position estimation to locate victims who need rescue promptly. To facilitate such rescuing tasks, quick and precise localization and tracking of police and medical staff for effective fleet management also becomes necessary. Various government and industry telecommunications committees have placed strict standards on the accuracy of positioning for emergency calls. Military applications has been one major force that sets up high standards and drives the research and development in this area. Another localization application is mobile gaming, whose performance heavily relies on the speed and preciseness of positioning as well. [1] and [11] provides an extensive overview on the history of source localization applications and techniques.

The localization application of primary interest here is microseismic event localization using a group of arbitrarily positioned passive sensors, which is critical to ensure the safety of underground construction and excavation activities. The specific application of interest that motivates this research is potash mining in Saskatchewan, Canada. Saskatchewan is "the 2nd largest potash producer" [12] and accounts for "40 percent of world potash trade" [13]. The immediate section below gives more detailed information about microseismic event localization using passive sensor arrays for potash mining in the Province of Saskatchewan.

1.2 Application of Interest

Potash, a salt ore that precipitated from an ancient sea that gradually evaporated [14], is of great value since it contains a fertilizer known as potassium chloride or KCL. KCL is one of three major fertilizers used by

farmers around the world, and Saskatchewan is blessed to have a huge deposit of potash salt underground. The ore, or the potash salt, is mined in a series of large, underground tunnels that may grow in any direction. According to [12], these mines are approximately one to three kilometers below the sedimentary rock depending on their geographical locations. The formation is shallower in southern Saskatchewan prairie compared to that in northern Saskatchewan Prairie. Due to the increasing world market demand for potash during the past more than half century, the mines in Saskatchewan had been expanding fast, "at a rate of approximately 500 meters/hour and at maturity, a mine can occupy an area of about 20 kilometers by 20 kilometers and have more than 6000 kilometers of tunnels" [14]. The expansion only slowed down recently due to the economic setback and big decrease in the price of KCL. However, it is expected that this situation is temporary, given the increasing population that will drive up the fertilizer demand in the long run.

A good general introduction to mining seismology can be found in [15]. Here some of the most basic concepts are briefly described. Microseismic events taking place in underground mines are usually induced earthquakes, resulting from mining operations that put surrounding earth material under high stress. When the tension of preexisting stress plus what is accumulated from mining activities reaches a certain level, the material may rupture or shift and there will come a sudden release of this tension. The energy will then propagate in the form of mechanical waves from the origin. These small induced earthquakes indicate potentially unstable areas where rock is likely to fall or burst. This is especially true for soft-rock mines, such as potash mines. A large potash mine can experience ten or even more microseismic events everyday, and there have been a number of fatalities reported because of rock falls. The aftermath of an accident due to rock rupture tends to become more devastating as mining operations go increasingly deeper into the earth to form a continuously growing network of underground tunnels. Therefore, monitoring the times and locations of microseismic events within the locale of a mine is extremely important to evaluate and prepare for mining hazards. Locating microseimic events plays a crucial role in maintaining safe mining activities underground. It can also help with the initial planning or ongoing expansion design of a mine, to avoid potentially hazardous areas.

The application of microseismic event localization in mines has much in common with those of seismology and geophysical exploration, though the latter focus on the study of the earthquake itself and the earth structure. In most cases, seismologists and geophysicists adopt the same type of sensors, usually geophones, to measure movement or vibration of the earth, because geophones are sensitive enough to respond to very distant and small ground motion. With the help of a seismic vibrator (most often a truck-mounted device), and/or some small-amount of dynamite, shock waves can be produced and travel in the earth. They are monitored by geophones deployed within the vicinity to yield a 3D map that can reveal the formation of gas, fluid and/or solid beneath the earth surface. Most oil, gas and mining companies often use this method to ensure that the mineral deposits of interest are present within the targeted area, so that the number of drilling pads required for site exploration and development can be greatly reduced. However, the concern in

this work is the location of low-intensity earthquakes in the locality of an underground mine, not the earth structure itself.

1.3 Problems Under Investigation

In this section, the three major problems studied in this work are described.

1.3.1 Preferred Algorithm

In contrast to many acoustic source signals studied in the literature whose duration can be several minutes or even longer, a microseismic event most often seen in an underground potash mine only lasts a fraction of a second [14]. Besides being transient, the emanated signal is a wideband signal made up of lower frequencies, instead of a narrowband signal of high frequencies [16]. These signal characteristics need to be considered when a source localization algorithm is being selected. This preferred algorithm should be unbiased, robust, and easy to implement.

Several representative unbiased estimators presented in the literature for source localization were chosen as candidates to initiate the work [17]. Then, simulations were done to evaluate the performance of each candidate to make a comparison among them. The final choice turned out to be the energy-based grid search (EGS) algorithm proposed in [14]. This preferred algorithm then served as part of the foundation for later research to be carried out, as mentioned in Section 1.3.2 and Section 1.3.3 below.

1.3.2 Cramér-Rao Bounds

Once a preferred algorithm is selected, it is desirable to know in theory how much room exists for performance improvement. A corresponding theoretical performance bound is needed to achieve this goal. One of the most widely used performance bounds is the Cramér-Rao Bound (CRB) [18], which expresses a lower bound on the variance of any unbiased estimator for an unknown deterministic or statistic parameter. An unbiased estimator asymptotically approaching the minimum variance bound is an efficient estimator. Due to its critical role in performance evaluation, CRB development has been of great interest to many researchers for decades and there have been many CRB expressions developed for various applications. Not surprisingly, one topic of this thesis is the CRB development for the evaluation of single-step acoustic source localization algorithms.

1.3.3 Optimization Techniques

Optimal signal processing is of great interest in various research areas, such as telecommunications, image compression, detection, estimation, and tracking. In both the time and frequency domain, there have been

good optimization techniques developed by different researchers. Hence, besides the CRB development, another major part of this thesis is to identify various optimization techniques that can be applied to the preferred algorithm selected earlier for short-duration signal localization to obtain better estimation accuracy.

1.4 Thesis Organization

The results for the study of optimal passive source localization outlined in Section 1.3 are presented in the rest of the thesis as follows.

1. Chapter 2 provides a literature review on some classical two-step and one-step source localization algorithms, CRB development, and optimization techniques available in various but similar application areas.

2. Chapter 3 presents a general system model that defines the source signal, the noise, the wave propagation, the sensors, and the sensor-source geometry adopted for the study in this thesis.

3. Chapter 4 studies and compares three chosen one-step and two-step source localization algorithms that are most representative in the literature. It reveals that the one-step EGS algorithm is preferred for the application of interest due to its robustness and simplicity.

4. Chapter 5 yields a few analytical expressions of the Cramér-Rao performance bound of one-step source location estimators for the single-path, multi-sensor scenario under consideration. Time continuous signals are considered instead of time discrete signals to make the work more consistent with that of [14]. Karhunen-Loève (K-L) expansion is used instead of the Fourier transform to make the mathematical derivation more general.

5. Chapter 6 proposes an amplitude weighting strategy to optimize the EGS algorithm. An analytical estimation error variance (EEV) expression of the amplitude weighted EGS algorithm (AWEGS) is developed. A closed-form optimal weight expression is then derived. A non-recursive method of applying the weights is proposed after it is mathematically proven to yield unbiased estimates. Monte Carlo simulation is performed to compare the AWEGS performance against the original EGS algorithm and the CRB developed in Chapter 5.

6. Chapter 7 proposes a frequency weighting strategy to optimize the EGS algorithm. An analytical EEV expression for the frequency weighted EGS algorithm (FWEGS) is developed. Closed-form optimal and suboptimal filter expressions are then derived, and their performances are evaluated and compared with one another using Monte Carlo simulations. The suboptimal filter expressions turn out to be classical filter expressions that are already given in the literature.

7. Chapter 8 presents and discusses all the results achieved, followed by an outline of some possible future

work.

References

[1] M. of the Underwater Acoustic Signal Processing (UASP) Technical Committee, "The past, present, and the future of underwater acoustic signal processing," *IEEE Sigal Processing Magazine*, vol. 15, pp. 21–51, August 1998.

[2] E. by Y.T. Chan, "Underwater acoustic data processing," NATO ASI Seriers, Kluwer Academic Publishers, 1989.

[3] J. Hassab, *Underwater Signal and Data Processing*. CRC Press, 1989.

[4] B. L. Daku, J. E. Salt, and C. M. McIntyre, "Quality of underwater source localization in a multipath environment," *J. Acoust. Soc. Am.*, vol. 91, pp. 957–964, February 1992.

[5] J. E. Salt and B. L. Daku, "Quality of multipath time-delay and source localization," tech. rep., Unversity of Saskatchewan, Saskatoon, Canada, 1991. Final Report for Defence Research Establishment Contract W7707-7-7933/01-OSC.

[6] A. Caiti, A. Garulli, F. Livide, and D. Prattichizzo, "Localization of autonomous underwater vehicles by floating acoustic buoys," *IEEE Journal of Oceanic Engineering*, vol. Vol.30, January 2005.

[7] M. Brandstein, J. Adcock, and H. Silverman, "A closed-form location estimator for use with room environment microphone arrays," *IEEE Trans. Speech, and Audio Processing*, vol. 5, no. 1, pp. 45–50, 1997.

[8] Y. Huang, J. Benesty, and G. Elko, "Passive acoustic source localization for video camera steering," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, pp. 902–912, 2000.

[9] J. Valin, F. Michaud, J. Rouat, and D. Letourneau, "Robust sound source localization using a microphone array on a mobile robot," in *IEEE/RSJ International Conf. on Intelligent Robots and Systems*, (Las Vegas, Nevada, USA), October 2003.

[10] R. Logeswaran, "Cost effective patient location monitoring system using webcams," *Journal of Medical Systems*, vol. 33(5), pp. 399–407, 2009.

[11] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: Possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Processing Magazine*, vol. 22, pp. 41–53, July 2005.

[12] "Potash introduction," 2017. Information on Website as of the writing of the final draft of this thesis.

[13] "Overview of potashcorp and its industry," 2014. Nutrients Section.

[14] B. L. Daku and E. J. Salt, "Error analysis of a localization algorithm for finite-duration events," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1024–1034, March 2007.

[15] S. J. Gibowicz and A. Kijko, *An Introduction to Mining Seismology.* Academic Press, 1994.

[16] A. Errington, B. L. Daku, D. Dodds, and A. Prugger, "Characterization of source signal in the potash mines," in *Proceedings of CCECE05*, (Saskatoon, SK, Canada), May 2005.

[17] S. Li and B. Daku, "Alternate amplitude weighting approach for passive source localization using the energy-based grid search algorithm," in *Proceedings of CCECE-2008*, CCECE at Sheraton Fallsview, Ontario, Canada, May 2008.

[18] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, 3rd ed., 1991.

2 Literature Review

This chapter summarizes acoustic source localization techniques developed mostly during the past half century that are the most relevant to this work.

The source signal position can usually be estimated by utilizing the time and spatial information contained within the signals captured by a set of passive sensors deployed at a certain distance away from the signal origin. Earlier work in this area mainly focused on narrowband sources in the far field, where sensors are distant from the source but close to each other and the sensors are usually organized in the form of a linear or circular array. The primary concern in this case is only the source signal's direction of arrival (DOA), not its precise geometrical coordinates.

Much of the recent research, including this work, focuses on wideband sources in the near field, where the distance between the sensors and the distance between the source and sensors are at about the same order. In this scenario, a good estimate of the source position is often desired. A wideband signal is defined here as one that has a ratio of the bandwidth to center frequency greater than one.

For mining applications, geophones are preferred for sensor network setup and the reasons will be discussed in more detail in Chapter 3. They are usually arranged along the mine rooms and are relatively arbitrary within the vicinity of where potential microseismic events may occur. This classifies these mining applications into the category of near-field problems.

The next two sections provide a brief review on the most representative acoustic source localization algorithms available in the literature for near-field applications. These algorithms, in general, can be categorized either as two-step algorithms or single-step algorithms. Then the third section provides some background information on optimization techniques for signal processing.

2.1 Two-Step Source Localization

To locate an acoustic source, an indirect two-step procedure presents a traditional method that first estimates the Time Difference of Arrival (TDOA) between any two sensors and then calculates the source location based on these TDOA estimates.

2.1.1 TDOA Estimation Techniques

Before any signal processing is performed, the sensors have to be organized in pairs, so that the relative time delays between each pair of sensors can be estimated using various techniques. All of these techniques then basically supply the signals received at the sensors as input to a signal processing system. This system yields a specific value or piece of data at its output, which is in correspondence with a certain TDOA estimate. Due to the presence of noise in the real world, there will always be some error in these TDOA estimates. There are three typical TDOA estimation techniques, which are briefly introduced below as examples.

Cross-Correlation Technique

The technique utilizing the cross-correlation function to estimate a TDOA has been well known and discussed in [1]. It is based on the characteristics of the cross-correlation between two signals with the same envelope but different arrival times. Theoretically, the value at which the cross-correlation function peaks represents the real TDOA value between the two signals. The cross-correlation function can be constructed in both the time and frequency domain.

Least Mean Square (LMS) Algorithm

In [2] It is implemented as an adaptive digital finite impulse response (FIR) filter that minimizes the mean square difference between its two inputs: a reference input, which drives the filter; and a desired input, which is a delayed version of the reference input. Its mean response is shown to converge to a discrete Wiener filter without a priori knowledge of the reference signal spectra [2]. The LMS algorithm has been widely applied in scenarios where the statistics of inputs are either unknown or at least partially unknown. The LMS algorithm can be implemented in either the time or frequency domain. The results are a set of weights at discrete times or frequencies which are used to estimate the time delay between the two inputs. More discussion on LMS algorithm can be found in [3], [4], [5].

Adaptive Eigenvalue Decomposition Algorithm (AEDA)

[6] proposes an algorithm for time delay estimation in an indoor environment, which is modeled as linear and time invariant. Therefore, the relation between a pair of sensor-received signals can be described with a linear system. This method focuses on the impulse responses between the source and microphones [6] in order to estimate the time delay. It consists of detecting the "direct paths of the two impulse responses between the source and sensors" [6], which are estimated in the eigenvector corresponding to the minimum eigenvalue of the covariance matrix of the sensor signals.

2.1.2 Localization Techniques using TDOA

With the TDOA estimates obtained using the techniques described in the previous section, there are traditionally two ways to further locate the source position: the hyperbolic intersection technique and the numerical search technique.

Hyperbolic Intersection

The hyperbolic intersection technique for navigating ships and aircraft has been in use for more than fifty years. Given the TDOA value estimated for each sensor pair, a set of non-linear equations that define a series of hyperbolic surfaces can be established and solved to get a solution that represents the source location estimate desired. When the signal propagation speed is a constant, the TDOA measurement constrains the source to a hyperboloid. The estimate of the source position is given by the intersection of these hyperboloids. The assumption originally adopted was that the location could be determined with only three receivers in two-dimensional scenarios. In 1972, Schmidt [7] recognized that four receivers are required to give unambiguous estimates for a two-dimensional location when using this hyperbolic technique, due to the need to account for the error contained within the TDOA estimates. Later, a number of suboptimal closed-form techniques have been developed based on spherical intersection [8], spherical interpolation [9], an improved estimation technique using two least square solutions [10], and a linear correction least squares estimator [11]. In [12], this widely adopted method is also applied to locating a mobile station.

Simplex Method

Besides the closed-form algorithms as mentioned immediately above, an example of using numerical search for source localization based on a set of TDOA values is given in [13], which uses the simplex algorithm to locate microseismic events in a mine. The simplex algorithm provides fast convergence and was originally described by Nelder and Mead in 1965 [14]. The basic idea is to utilize an $L+1$-dimensional general simplex in a L-dimensional space to minimize a target error function that provides vertex values for the simplex. The target error function is defined as the difference between the calculated and estimated values of a (group of) certain parameter(s) of interest. The entire region of the misfit errors is defined [14] as the "error space". After the error values at the $L+1$ vertices of the simplex are compared with each other, the vertex with the highest error value is replaced by some other point in the L-dimensional space with a lower error value. As the procedure is repeated, the shape of the simplex moves and distorts to converge to the best-fit solution. For mining applications, each set of coordinates in the L-dimensional "error space" represents a possible location of the microseismic event to be estimated and is related to a value of the error function for the TDOAs. The error is calculated by comparing each TDOA value corresponding to a hypothesized source position with the observed real world TDOA value. If there are K sensor pairs given K TDOA values, a K-variable misfit error function can be formed and then minimized to locate the source position utilizing

such a simplex in the corresponding "error space".

2.2 One-Step Source Localization

The performance of two-step localization algorithms rely heavily on the accuracy of TDOA estimates. When there are large errors existing in the TDOA estimates, the source location estimate is prone to even larger error due to its non-linear relationship to the TDOA. Many applications that use TDOA estimates, for example, underwater sonar, usually involve a very long duration signal. This signal is then truncated for further processing, but with a length long enough to produce a large time-bandwidth product so that the system can perform above threshold [15]. Such two-step indirect methods do not work as well for signals with small time-bandwidth products and low signal-to-noise ratios (SNR) though, such as those encountered frequently in an underground mine. As a consequence, other techniques that can avoid the TDOA estimation procedure are desired.

One alternative is a direct procedure named the Approximated Maximum Likelihood (AML) algorithm proposed in [16], which is implemented in the frequency domain. It approaches the problem using parametric array processing to fully exploit the underlying data model for good efficiency and robustness, at the cost of extensive multidimensional search and signal processing.

Another alternative single-step estimator is also a search-based algorithm that maximizes a function over a grid of hypothesized source locations. It is proposed in [17] with the name Energy-based Grid Search (EGS), which directly uses the original time-domain signals received at the sensors.

Examples of previous work related in this area include the performance analysis of a multipath, multi-sensor underwater acoustic algorithm using a summed correlator [18], and a speaker localization problem addressed using a summed correlator algorithm, referred to as a steered filter-and-sum beamformer, in [19,20].

The two single-step source position estimators that do not require explicit TDOA estimates, AML and EGS, are of the most interest and farther described below.

Maximum Likelihood Estimation

The well-known and frequently used Maximum Likelihood (ML) method is model-based and requires a statistical data framework. In many cases, the total noise at each sensor can be modeled as a white Gaussian random process with a zero-mean and common variance, whereas the signal waveform is treated as deterministic (arbitrary) and unknown. Then the conditioned likelihood function can be easily determined with the aid of the probability density function (p.d.f.) of all the observations for the unknown parameter(s), in either the time or frequency domain. The desired parameter to be estimated under discussion is the source location, which can then be obtained by solving a minimization problem numerically corresponding to this

conditioned likelihood function. Given a good initial guess, this method usually converges quickly to a minimum.

A parametric ML approach to solve for a wideband source location in a near-field scenario is proposed in [16]. This solution, named as the Approximate Maximum Likelihood (AML) algorithm in [16], is developed based on the classical maximum-likelihood DOA estimator for narrowband signals. The received wideband signal is first transformed using the Discrete Fourier Transform (DFT) into the frequency domain, where the signal spectrum is represented by the narrowband model for each frequency segment. The AML method is then applied to each positive frequency segment to estimate the source location, and a combination of the AML metrics for all the frequency segments yields the final location estimate. In addition to the grid search, a large number of samples is required to perform the DFT well enough. The AML algorithm is hence calculation intensive. In cases where there is no a prior information to narrow down the size of the possible region containing the source location, the calculation load becomes even heavier. To reduce the number of points involved in the calculation required, the likelihood function can be first evaluated only on carefully selected points on a coarse non-uniform grid that covers a large area. For example, polar coordinates with non-uniform sampling of the range and uniform sampling of the angle can be used. Also the grid points can be arranged denser near the array and sparser away from the array. After a crude estimate of the source location is obtained from one or more coarse grid-point searches, finer grids can be applied within its vicinity to reach the global minima more accurately.

Energy-based Grid Search Algorithm

The other single-step estimator of interest is an algorithm analyzed in [17] and named Energy-based Grid Search (EGS) in the context of this study. This EGS algorithm, like the AML, is based upon a global search of a three-dimensional grid. The grid limits the search to a box of finite dimensions. The grid can be non-uniform if needed. The EGS algorithm estimates the source location by finding within a given space the hypothesized position that produces the greatest energy in the sum of all the sensor signals after each of which has been time-shifted by an amount consistent with the propagation delay of the hypothesized location of the event. The theoretical bases of this estimator is the cross-correlation and auto-correlation. The total energy function used is basically the sum of all the cross-correlations and auto-correlations of the sensor signals. When the sensor signals with the same envelope but different time lags are aligned in time, the sum of cross-correlations and auto-correlations peaks. This brings an algorithm that works very robustly with short-time duration signals of a small time-bandwidth product [17]. Due to the presence of random noise, the location of the prominent peak of the total energy function moves randomly, thus creating estimation error in practice. One major difference between the EGS and AML algorithm is the domain in which most data processing takes place. The former is in the time domain while the latter is in the frequency domain.

An analytic expression for the estimation error variance of the EGS localization algorithm is developed

in [17] and forms the foundation of this study. Both the analysis in [17] and this work utilizes the prominent peak in the total energy function, which shifts around the actual position of the source in the presence of noise. Some underlying assumptions made are that the source signal is finite and it has no discontinuity, which are acceptable for the given application scenario. It is also assumed that the mesh used for the grid is fine enough or that high order interpolation is used so that the position of the actual peak can be located with desired precision.

2.3 Cramér-Rao Bounds

For a certain application scenario that can be defined with a set of known and unknown parameters, it is always desirable to understand what the best estimation accuracy can be achieved in theory for a certain unknown parameter to be estimated, regardless of which unbiased estimators are used. Such a theoretical performance bound is also very useful in evaluating algorithms for a specific application environment of interest. The most widely used performance indicator is the Cramér-Rao Bound (CRB), which provides a lower bound expression on the variance of any unbiased estimators, for a deterministic or statistic parameter [21].

Some examples of work on CRB include CRB expressions developed for estimating the channel and signal parameters, or the signal itself in additive White Gaussian Noise (WGN) scenarios [22–24]. In [25] and [26], CRB expressions are presented for estimating discrete frequency components in the presence of special Coloured Gaussian Noise (CGN) that can be represented analytically by functions of WGN. Though providing good insight into the CRB development procedure, the above references are not directly related to the application of interest. As for source localization problems, most CRB expressions developed are only in terms of TDOA or DOA estimation errors, such as in [27–29]. These are not helpful for this study either, which is interested in the source location estimation accuracy.

[17] suggests that when the time-bandwidth product of a source signal is small, the one-step estimator that avoids the TDOA estimation procedure is more robust than the two-step estimators that rely on TDOA estimates. (The unit of the time-bandwidth product is in $second \cdot hz$). Hence a CRB expression is desired to provide a performance indicator that can help evaluate all the single-step source location estimators for the scenario briefly described in 2.2.

[12] gives a CRB expression for the final location estimation error that falls in our area of interest, but it is derived on the basis of assumptions that the error variance introduced by previous TDOA estimates is already given. This makes it applicable only to cases where long-signal records are available such that the TDOA can be estimated with acceptable error. [16] focuses on the source localization application using a single-step source location estimator, but the corresponding CRB is derived in such a way that the sensor signal amplitude is independent of the distance between the source and sensors. This is not appropriate for

most near-field applications. Though [16] states it considers the case where the source position and source signal samples are both unknown, the contribution of unknown source signal samples to the CRB for the location estimation error is actually ignored during its derivation. In addition, [16] only considers WGN and discrete time signals. In this study it is of interest to start with the original continuous signals before sampling and to consider both WGN and CGN, in order to make the derivation presented here as general as possible.

2.4 Optimization Techniques

This section provides a brief overview of some optimization techniques for signal processing developed in the literature that bring insight to this study.

2.4.1 Beamforming

Most of the work devoted to optimal array signal processing has been for far-field applications focused only on the DOA or TDOA estimation. A widely adopted technique for this is beamforming, which is implemented with the aid of a linear or circular antenna array whose sensors are usually evenly spaced [30–32].

In these works, the source signal is normally assumed narrowband modulated on a high frequency carrier. The goal of beamforming is to have the power output of a sensor system maximized after all the sensors are managed to steer towards the direction of signal arrival, so that more accurate DOA/TDOA estimation can be achieved. The signal amplitude received across the array is treated equal and this does not significantly affect the estimation result for far-field cases. Hence the optimal steering vector is determined only by the DOA and distance between receivers.

In near-field applications, the received signal amplitude being different at each sensor cannot be ignored and should be utilized to locate the source position. Following the line of thoughts of the beamforming technique for far-field scenarios, one task of this work is the maximization of the sensor array power output, after non-uniform sensor signal amplitudes are accommodated in the algorithm.

2.4.2 Frequency Domain Techniques

The beamforming techniques introduced above are mostly rooted in the time and space domain. With the help of the Fourier transform, optimization techniques in the frequency domain are also considered in this work. Actually, optimal filter design has been widely studied to improve system performance in many applications, and most filter design techniques are based firmly on frequency domain concepts. Both one-step and two-step source localization algorithms can benefit from properly designed filters. In this section, filters that are optimal in terms of various performance indices for various applications in the literature are briefly

introduced.

Rough Emphasis on Certain Frequency Bins

A sensor array with predefined spatial and frequency constraints is discussed in [33] and [34]. A wideband signal is considered and an evenly spaced linear sensor array is constructed so that "maximum energy concentration can be achieved over a desired spatial 'look' and some frequency regions" [34], subject to certain conditions given in the scenario considered.

After the spatial spectral density matrix is obtained, the signal subspace property is exploited. Based on the Wideband Multiple Signal Classification (MUSIC) algorithm, correct steering vectors that are orthogonal to the null subspace are found to estimate a single or multiple source locations in [35].

Filters for Optimal Time Delay Estimation

[36] and [37] describe optimal signal processing for passive range and bearing estimation using the traditional two-step process. Their common focus is on filter design applied during the first step: to optimize the accuracy of time delay estimation. This is achieved by minimizing the frequency domain ML estimation error of the TDOAs with the aid of a separate correlator delay measurement system. Incorporating the sensor geometry, signal spectra and noise spectra, linear combinations of these delays can be weighted to yield better source range and bearing estimates.

Another means to optimize the TDOA estimation is to prefilter the signals before computing their correlation. This gives the so-called weighted Generalized Cross-Correlation (GCC) method in the frequency domain that is well discussed in [38]. The expression of the weighted GCC for two received signals consists of the crosspower-spectrum between them and a prefiltering function. Identifying the expression for this prefiltering function is crucial, since by emphasizing specific spectral components, it can affect the TDOA estimation greatly.

In [1], several frequency weighting strategies are discussed. When there is a priori information about the source signal characteristics available and the noise can be assumed white Gaussian, the peak of cross-correlation that brings the TDOA estimate will be sharpen by whitening the input signals, which can be done very conveniently. A deterministic approach is presented in [1], while a statistical approach to the signal whitening problem is proposed in [38].

In [39], a weighting technique is introduced as the so-called Phase Transform (PHAT), which works independently of input signal waveforms. This characteristic enables PHAT to work well for wideband signals with no a priori information available. It is a different story for narrowband signals though, since PHAT amplifies the background noise in this case.

Filters for Minimal Estimation Error Variance

The estimation error variance (EEV) is the most suitable performance index for an estimator if an analytic expression for the EEV is achievable. None of the above filters was derived to be optimal in terms of the EEV. Hence in the context of this work, a filter that can minimize the EEV value is defined as the "optimal filter".

In [40] and [41], the application of registering a received image (with noise) with a known ideal image is considered. The authors successfully derived a closed-form expression for the image registration error variance. Then, they proved that the matched filter is the optimal filter that minimizes this EEV value under the assumption of additive WGN.

The matched filter has been widely known and discussed extensively in most communications textbooks. Generally speaking it does not refer to any specific filter structure, but defines what a filter achieves: maximizing the system SNR in consideration. In most cases where noises are assumed to be additive and independent WGN, the matched filter takes the form of the original transmitted signal after it is delayed and transposed. Using a spectrum representation, the transfer function for the matched filter is the complex conjugate of the Fourier transform of the original source signal with a delay constant for all the frequency components.

This work shows that the matched filter is not always the "optimal" one for a certain parameter estimator of interest.

References

[1] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 24, pp. 320–327, August 1976.

[2] Y. T. Chan, R. V. Hattin, and J.B.Plant, "The least squares estimation of time delay and its use in signal detection," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. Vol.ASSP-26, June 1978.

[3] F. A. Reed, P. L. Feintuch, and N. J. Bershad, "Time delay estimation using the lms adaptive filter-static behavior," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-29, pp. 561–571, June 1981.

[4] D.H.Youn, N. Ahmed, and C. Carter, "On using the lms algorithm for time delay estimation," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. ASSP-30, pp. 798–801, October 1982.

[5] T.L.Tung, K.Yao, C. Reed, R.E.Hudson, D.Chen, and J.Chen, "Source localization and time delay estimation using constrained least squares and best path smoothing," in *Part of the SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations IX*, vol. 3807, (Denver, Colorado), SPIE, July 1999.

[6] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive," *Acoustical Society of America.*, vol. 107, pp. 384–391, January 2000.

[7] R.O.Schmidt, "A new approach to geometry of range difference location," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-8, pp. 821–835, 1972.

[8] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 35, pp. 1223–1225, August 1987.

[9] J. O. Smith and J. S. Abel, "The spherical interpolation method of source localization," *IEEE J. of Oceani. Eng.*, vol. 12, pp. 246–252, January 1987.

[10] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. on Signal Processing*, vol. 42, pp. 1905–1915, August 1994.

[11] Y. Huang, J. Benesty, G. Elko, and R. Mersereau, "Real-time passive source localization: a practical linear-correction least-squares approach," *IEEE Trans. on Speech and Audio Processing*, vol. 9, pp. 943–956, November 2001.

[12] K.W.Cheung and H.C.So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE Trans. on Signal Processing*, vol. 53, pp. 460–470, February 2005.

[13] A. F. Prugger and D. J. Gendzill, "Microearthquake location: a nonlinear approach that makes use of a simplex stepping procedure," *Bulletin of th Seismological Society of America*, vol. 78, pp. 799–815, April 1988.

[14] J. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.

[15] J. P. Ianniello, "Large and small error performance limits for multipath time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 245–251, April 1986.

[16] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. on Signal Processing*, vol. 50, pp. 1843–1854, August 2002.

[17] B. L. Daku and E. J. Salt, "Error analysis of a localization algorithm for finite-duration events," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1024–1034, March 2007.

[18] S. W. Davies and M. A. Price, "Source localization by summing multiple correlator outputs," in *Proceedings of the ICASSP 90*, vol. 5, pp. 2787–2790, 1990.

[19] N. Strobel, T. Meier, and R. Rabenstein, "Speaker localization using a steered filter-and-sum beamformer," in *Erlangen Workshop: Vision, Modeling, and Visualization*, (Erlangen, Germany), November 1999.

[20] N. Strobel and R. Rabenstein, "Robust speaker localization using a microphone array," in *Proceedings of the European Signal Processing Conference (EURASIP)*, vol. III, pp. 1409–1412, 2000.

[21] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, 3rd ed., 1991.

[22] F. Gini, M. Luise, and R. Reggiannini, "Cramér-Rao bounds in the parametric estimation of fading radiotransmission channels," *IEEE Trans. on communications*, vol. 46, pp. 1390–1398, October 1998.

[23] F. Gini and G. B. Giannakis, "Hybrid FM-polynomial phase signal modeling: Parameter estimation and Cramér-Rao bounds," *IEEE Trans. on Signal Processing*, vol. 47, pp. 363–377, February 1999.

[24] M. Ghogho, A. K. Nandi, and A. Swami, "Cramér-Rao bounds and maximum likelihood estimation for random amplitude phase-modulated signals," *IEEE Trans. on Signal Processing*, vol. 47, pp. 2905–2916, November 1999.

[25] J. M. Francos and B. Friedlander, "Bounds for estimation of complex exponentials in unknown colored noise," *IEEE Trans. on Signal Processing*, vol. 43, pp. 2176–2185, September 1995.

[26] D. N. Swingler, "Approximate bounds on frequency estimates for short cisoids in colored noise," *IEEE Trans. on Signal Processing*, vol. 46, pp. 1456–1458, May 1998.

[27] H. Ye and R. D. Degroat, "Maximum likelihood DOA estimation and asymptotic Cramér-Rao bounds for additive unknown colored noise," *IEEE Trans. on Signal Processing*, vol. 43, pp. 938–949, April 1996.

[28] J. P. Delmas and H. Abeida, "Stochastic Cramér-Rao bound for noncircular signals with application to DOA estimation," *IEEE Trans. on Signal Processing*, vol. 52, pp. 3192–3199, November 2004.

[29] H. Abeida and J.-P. Delmas, "Gaussian Cramér-Rao bound for direction estimation of noncircular signals in unknown noise fields," *IEEE Trans. on Signal Processing*, vol. 53, pp. 4610–4618, December 2005.

[30] S. P. Applebaum, "Adaptive arrays," *IEEE Trans. Antennas and Propagation*, vol. 24, pp. 585–598, September 1976.

[31] W. F. Gabriel, "Spectral analysis and adaptive array superresolution techniques," in *Proceedings of the IEEE*, vol. 68 of *6*, pp. 654–666, June 1980.

[32] G. Carter, "Time delay estimation," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 29, June 1981. Special Issue.

[33] T. Tung, K. Yao, D. Chen, R. Hudson, and C. Reed, "Source localization and spatial filtering using wideband MUSIC and maximum power beamforming for multimedia applications," in *1999 IEEE Workshop on Signal Processing Systems: Design and Implementation*, (Taipei, Taiwan), pp. 625–634, SIPS 99, 1999.

[34] D. Korompis, K. Yao, and F. Lorenzelli, "Broadband maximum energy array with user imposed spatial and frequency constraints," in *Proceedings of ICASSP-94*, vol. iv, pp. IV529–IV532, ICASSP-94, April 1994.

[35] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood and Cramér-Rao bound: Further results and comparisons," *IEEE. Trans. on ASSP*, vol. 38, pp. 2140–2150, December 1990.

[36] W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. on Information Theory*, vol. IT-19, pp. 508–614, September 1973.

[37] W. R. Hahn, "Optimum processing for passive sonar range and bearing estimation," *J. Acoust. Soc. Am.*, vol. 58, pp. 201–207, July 1975.

[38] W.K.Pratt, "Correleation techniques for image registration," *IEEE Trans. on Aerosp. Electron. Syst.*, vol. AES-10, pp. 353–358, May 1974.

[39] S. Alliney and C. Morandi, "Digital image registration using projections," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAPI-8, pp. 222–233, March 1986.

[40] C. D. Mcgillem and M. Svedlow, "Optimum filter for minimization of image registration error variance," *IEEE Trans. on Geoscience Electronics*, vol. GE-15, pp. 257–259, October 1977.

[41] C. D. Mcgillem and M. Svedlow, "Image registration error variance as a measure of overlay quality," *IEEE Trans. on Geoscience Electronics*, vol. GE-14, pp. 44–49, January 1976.

3 System Model

In order for the topics presented in Chapter 1 to be investigated, the following basic elements must first be specified: the source signal, the way the signal is propagated, the noise, the sensors, and the geometry of the sensor system adopted for data acquisition in this thesis. It is also very important to identify which parameters are known and which are unknown in the problems under investigation, since they are critical for the performance bound to be derived correctly and for the proposed theoretical algorithms to be implemented effectively. Throughout the research work conducted here, the propagation speed and sensor positions are taken as known parameters. The location of the source signal is assumed to be an unknown parameter; the noises are assumed to be parameters that are statistically known by their first and second moments.

3.1 Source signal

In this work, the source signal is defined to be one single-point acoustic signal with a low time-bandwidth product, originating from a microseismic event.

Microseismic events are a type of earthquakes, which can be caused by either natural occurrences or human triggers. There are various reasons humans "create" small scale earthquakes: for scientific and experimental purposes or for industrial and commercial purposes. The specific trigger of an human induced earthquake can be many, such as a dynamite explosion, compressed air gun shot, vibrating device, and/or regular daily routine mining activity. Most of the time, the induced earthquakes from these man-made causes are imperceptible to humans a certain distance away from the source, and therefore referred to as microseismic events.

The Richter magnitude scale was widely used for a long time to classify earthquake sizes quantitatively. Now the moment magnitude scale is used, since it provides better resolution [1]. In the context of this thesis, an earthquake with a moment magnitude less than 1 is referred to as a microseismic event, or microearthquake. In the potash mines of Saskatchewan, Canada, a microseismic event typically generates only 750 to 1000 joules of mechanical wave energy [2], which is far less than that of a 1-moment-magnitude earthquake.

The microseismic event source in consideration is defined as a single point or a monopole source. It is abstracted as a particle as opposed to other more complex types, such as dipoles or quadrupole sources, because the possibility of more than one microseismic event taking place side by side at the same time is extremely low, and the radial distance in concern is much larger than the dimension of the vibration source generating the acoustic waves. Throughout this study it is assumed that within a given short period of time, only one microseismic event takes place, which is reasonable considering how short a microseismic signal usually lasts. Multiple-source localization is therefore not of concern within the scope of this research.

"Acoustic" in the Merriam-Webster dictionary is defined as "of or relating to the sense or organs of hearing, to sound, or to the science of sounds". Indeed, it is a broad term that touches on almost the whole realm of mechanical vibration and waves, whether within the audible spectrum or not, and regardless of the medium in which the signals travel. For the specific type of microseismic event signals of interest here, the frequencies are observed to be low and concentrating roughly within the range of 50 - 300 Hz [3] and the medium is solid rock. Since the time duration of the signal is very short, which is only a fraction of a second, the time-bandwidth product of the signal is usually less than 50 and considered small.

For a given microseismic event, the source signal can be viewed as deterministic and unknown, or it can be viewed as one ensemble of a random process. Some statistics of such a random process may already be known from past events observed, since all the microseismic events induced by the same type of phenomenon or activity within the same type of medium bear similar properties. The majority of this thesis assumes the source signal to be deterministic and unknown. Chapter 7 though, investigates the situation where the source signal is treated as partly known, since its frequency information is statistically available.

3.2 Wave Propagation

Another important element requiring a clear definition in the system model is the wave propagation of a seismic event.

When an earthquake occurs, transverse and/or longitudinal mechanical waves that originate from the event source carry the energy burst out and usually propagate outwards omni-directionally. These seismic waves are able to spread out because the earth material possesses a certain degree of elasticity to transmit them.

Regular mining activities can cause shear rupture in rocks and puts stress, including compression, tension and/or shearing, to the surrounding earth material. This, as a consequence, makes the surrounding earth material undergo strain, which causes deformation in either the volume or shape. The elasticity of a material is reflected in its strain rate, defined by the length of time the material takes to reach a certain level of deformation under a given amount of stress. Material with good elasticity recovers from its deformation

and goes back to its original volume and shape when the stress ceases. The general Hook's law describes a linear relationship between the stress and strain in an elastic medium, which is true when the amount of stress applied to the medium is within a certain range. Beyond this range the material will pass its elasticity limit and will either respond with brittle fracturing, which is the main cause of sudden falls and bursts in the rock, or ductile/plasticity behavior, meaning the deformation stays in place even when there is no more stress present. However, the inelastic behaviour of seismic waves only need to be taken into account for either very large-amplitude seismic-caused deformations in soft soil, or for extremely long-period of free-oscillation, neither of which is the case here. More detailed information about seismic wave propagation and earth models can be found in [4].

In this study, the medium in which seismic waves travel is solid rock and assumed isotropic, which means waves radiate outwards from the origin uniformly and omni-directionally. In this case, "the number of independent parameters in the elastic tensor described by the Hook's law is reduced to just 2 from 21 [4]." If the properties of the medium vary when measured along the axes in different directions, the medium is called anisotropic and the wave propagation will be direction-dependent. Although "in some parts of the earth's interior, anisotropy on the order of a few percent exists, isotropy has proven to be a reasonable first-order approximation for the earth as a whole" [5].

Seismic waves are mainly defined as body waves or surface waves, based on their travel paths. As their names suggest, body waves spread through the lithosphere, the outer part of the earth's body, while surface waves only find their way along the surface of the planet, somewhat similar to wavelets often seen on water.

Body waves include P-waves and S-waves, where P stands for primary and S for shear or secondary. The name "primary" results from the fact that P-waves propagate the fastest among all kinds of seismic waves generated during a earthquake and hence the first being captured by a seismometer. Similarly, the name "secondary" is derived from the fact that S-waves are in general the second one to be recorded by a seismometer. Another name for P-waves is "compressional waves", because they cause the medium through which they travel to change volume from being pressed to stretched alternatively by pushing and pulling the medium particles in the same direction the waves propagate. Depending on the medium being solid, thick liquid, water-like fluid or gas and the direction of the wave propagation, the P-wave speed varies. In comparison, S-waves can only travel through the solid at about 60% of the speed of P-waves in the same type of medium (Wikipedia: Seismic waves). S-waves cause solid particles to have movements that are vertical to the wave propagation direction. Despite what their names suggest, earthquakes generate much stronger secondary waves than primary waves according to historical data records. This has also proven to be true for microseismic events of interest here, since "P-waves are rarely observed on the monitoring systems used in Saskatchewan potash mines" [6].

Surface waves are the result of body waves interacting with the surface and lithosphere of the earth. It is

obvious that major damage done in an earthquake is by surface waves, since man-made infrastructures are mostly built on/beneath the surface of the earth. The dreadful ground shaking a human body can sense is mostly caused by surface waves as well.

There are also two subcategories under the name "surface waves": the Love waves and Rayleigh waves. No introduction is needed for them here, because the microseismic events of interest that can happen a dozen times a day are not strong enough to produce any such surface waves. Hence they are not part of the system model that needs to be considered for this study. [7] provides more detailed description for various types of seismic waves and their characteristics. For different mediums, the mechanical wave characteristics and behavior can be modeled similarly, but the wave speed and attenuation rate are different depending on the elastic and inertial factors of the mediums. "Typical S-wave propagation speeds are on the order of 1 to 8 km/sec. The lower value corresponds to the wave speed in loose, unconsolidated sediment, the higher value is near the base of the earth's mantle [7]." In the application of interest, the S wave velocity observed in a potash mine is roughly 2500 m/s [2].

A wave traveling in a heterogeneous material may undergo refraction, reflection, dispersion, diffraction, and attenuation. Please see [8] for a more in-depth description of seismic wave propagation in heterogeneous media. However, the medium of the propagation model adopted for this study is assumed homogeneous, which then excludes refraction, dispersion, and diffraction. This assumption is reasonable since within the vicinity of a mine usually the subsurface geological characteristics remain the same. Therefore, the only type of interaction that requires consideration is reflection. Paths involving reflection are longer, which means the reflected signals are delayed relative to the signal traveling directly from the source to sensor in the same type of medium. The strengths of the delayed reflected signals are thus usually significantly weaker than the direct path signal, and it is reasonable to ignore the signals arriving other than those along the direct path. Therefore, a single path model is suitable to address the problem of interest.

It is also assumed that the wave amplitude decreases inversely proportional to the distance it travels from the origin and that no energy is lost. Therefore, a system model assuming that a source signal spreads spherically with a constant propagation speed can be established, provided that the medium considered is isotropic and homogeneous.

3.3 Sensors

The sensing components considered for the system model in this study is a type of seismometers that have been widely used: the geophones. These electromechanical devices are constructed based on the physics stated in Newton's Law of Inertia: an object tends to remain in whatever state it is currently in, until an external force causes it to change.

Microseismic events generate signals that last a short period of time and only cause the earth within the vicinity to experience limited shaking. The geophones are selected for the application of interest because they satisfy the desired requirements for sensitivity and operational frequency range. They produce output signals proportional to velocity instead of displacement or acceleration, with a relatively high sensitivity. Their cost is relatively low due to the simplicity in their structure.

A typical geophone is made up of a magnetic component, a cylinder wrapped by a copper wire coil, and a spring. The cylinder hangs from the spring that is attached to the magnetic housing and the housing is fixed to the earth. When the earth is shaking, the magnetic housing moves along but the cylinder tends to stay where it is. The relative movement between the copper wire coil on the cylinder and the magnet forms an electromagnetic field and creates current. The current is then output as voltage, which is proportional to the velocity of the ground movement, representing the microseismic wave propagation speed.

The geophones with a structure as described above are passive analog devices made up of electromechanical components. Technology advancement has brought into the world a new kind of seismometers that are made up of microelectromechanical components. Unlike the geophones that track the earth movement, these new products utilize an active control loop to help a small piece of silicon resist the earth movement and stay in its original position. By doing so, these devices generate responses proportional to acceleration, which categorizes them as accelerometers.

Most of the cases, microelectromechanical accelerometers provide wider responsive frequency ranges compared to geophones. However, the benefit comes at the cost of suffering a much higher noise level, due to the nature of its structure. The existence of thermal/shot/$\frac{1}{f}$/generation-recombination noise has already been widely and well observed in microelectronic devices. The use of microcantilevers or membranes to manufacture a microelectromechanical device makes the situation even worse by introducing special extra noises. The geometrical prints of these components are so small that they tend to cause problems during the heat dissipative process. This then causes temperature fluctuation within various vibrational structures of such accelerometers, which results in additional noises specific to microelectromechanical devices. Therefore, though these devices may make a great candidate for strong vibration applications with long lasting signals, they are generally not suitable for microseismic events that are often seen in potash mines.

When dimension is not a concern, traditional geophones with a simple-structure are almost still the best choice among all seismometers, considering its lower cost with good linearity, dynamic range and lower noise generation. Modern microelectromechanical accelerometers do not compete well for applications involving weak and transient signals like microseismic waves.

Readers can find more information about seismometers and geophones in [9].

Given the simple structure, a classical geophone is only able to measure a single-dimensional movement of

the earth, although in reality seismic pressure waves shake the ground in three dimensions. This can presents a problem when the full wave needs to be observed and utilized. However, P-waves are rarely observed in Saskatchewan potash mines [6] and the concern here is only the location of a microseismic event, not the waves. It is therefore acceptable to adopt and place these single-dimensional geophones on the ceiling of the mining tunnels to monitor S-waves oscillating along one direction.

For the application of interest, it is assumed that an array of sensors are used to record acoustic pressure waves produced by a microseismic event. Each sensor includes a classical geophone and a data transmission system. All the sensors adopted bear the same mechanical and electrical properties.

3.4 Noise

It is reasonable to assume as in [2] that the output of each sensor, which consists of a geophone and a data transmission system, contains noise that is a combination of ambient vibrations and Johnson noise. It is acceptable to model the noise as white Gaussian in this case. However, sometimes a stationary coloured Gaussian noise (CGN) model may better reflect the physical situation, due to the existence of electronic elements. The spectrum of the actual source noise will be shaped by these elements as the noise passes through the sensor before entering a signal processing device for source location estimation. Also in some cases, there may be a certain interfering signal that can be either known or characterized as a Gaussian process. Therefore, for the investigation covered in this work, either a WGN or stationary CGN model is adopted for the noise, depending on the scenario of a specific research topic under discussion. However, regardless of whether the type of noise is WGN or CGN, the noise at all sensors is assumed to be additive and spatially independent, and to have an identical power spectra. To be more precise, the noise is assumed to have identical distribution with a zero-mean and common variance.

3.5 Sensor System

In the sensor system adopted for the application of interest, it is assumed that the response of every sensor is identical,and the output of each sensor is continuous and contains Gaussian noise as described in Section 3.4.

The spatial coordinate system used here is the three-dimensional x, y, z Cartesian system, similar to the one used in [2]. The source position is given by $\vec{p}_s = (x_s, y_s, z_s)$. There are M sensors placed arbitrarily and sparsely in the vicinity of the source. The position of sensor m is denoted $\vec{p}_m = (x_m, y_m, z_m)$. The distance from the source to sensor m is denoted d_m and is given by

$$\begin{aligned} d_m &= \| \vec{p}_m - \vec{p}_s \| \\ &= \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2 + (z_m - z_s)^2}. \end{aligned} \tag{3.1}$$

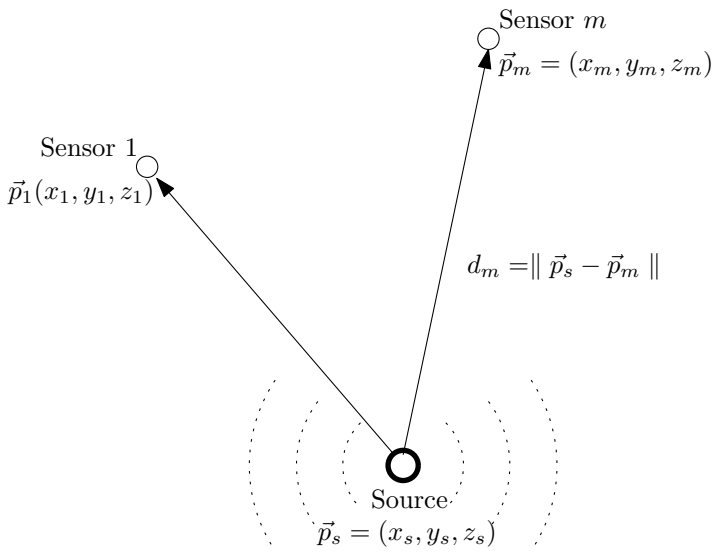

Figure 3.1: Coordinate system for the sensors

The output of sensor m, denoted $r_m(t)$, is a continuous sample function of a stochastic process $\mathbf{r}_m(t)$ that is given by

$$\mathbf{r}_m(t) \quad = \quad s_m(t) + \mathbf{n}_m(t), \tag{3.2}$$

where $\mathbf{n}_m(t)$ is an additive noise random process, and $s_m(t)$ is the signal received at sensor m that is independent of the noise. $s_m(t)$ is simply a delayed and attenuated version of the original deterministic source signal $s_0(t)$ and is given by

$$s_m(t) \quad = \quad \frac{s_0(t - \tau_m)}{d_m}, \tag{3.3}$$

where $\tau_m = \frac{d_m}{c}$ is the time for the original signal $s_0(t)$ to travel from the source to sensor m, with c being the propagation speed of the mechanical wave. To accommodate this propagation model, it is assumed that all the observed waveforms are collected by sensors that are at least one unit distance away from the source. The geometry of the sensor system is shown in Figure 3.1.

It is reasonable to assume that $s_0(t)$, denoting the real source signal, contains no discontinuity and is finite with a short-time duration T. The start time of the source event is taken as the time origin, and T_o is the end of the observation time, which is assumed much greater than T. Mathematically, T_o may extend to infinity.

References

[1] W. Spence, S. A. Sipkin, and G. L. Choy, "Measuring the size of an earthquake," *Earthquakes and Volcanoes*, vol. 21, no. 1, 1989.

[2] B. L. Daku and E. J. Salt, "Error analysis of a localization algorithm for finite-duration events," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1024–1034, March 2007.

[3] A. Errington, B. L. Daku, D. Dodds, and A. Prugger, "Characterization of source signal in the potash mines," in *Proceedings of CCECE05*, (Saskatoon, SK, Canada), May 2005.

[4] P. Bormann, B. Engdahl, and R. Kind, *Seismic Wave Propagation and Earth models*. GFZ Helmholtz-Zentrum Potsdam, 2012.

[5] P. Bormann, B. Engdahl, and R. Kind, *New Manual of Seismological Observatory Practice (NMSOP)*. Potsdam : Deutsches GeoForschungsZentrum GFZ, 2009.

[6] A. Errington, *Sensor Placement for Microseismic Event Location*. M.sc thesis, University of Saskatchewan, 2002.

[7] C. J. Ammon, "Slu eas-a193 class notes on seismic waves and earth's interior," tech. rep., Department of Geosciences, Pennsylvania State University, the United States.

[8] G. D. Manolis, P. S. Dineva, T. V. Rangelov, and F. Wuttke, *Seismic Wave Propagation in Non-Homogeneous Elastic Media by Boundary Elements*. Springer International Publishing, 2017.

[9] C. J. Ammon, "Slu eas-a193 class notes on waves, seismograms, and seismometers," tech. rep., Department of Geosciences, Pennsylvania State University, the United States.

4 Performance Comparison for Three Source Localization Algorithms

The content was originally published as:

S. Li and B. L. Daku, "Performance Comparison for Three Source Localization Algorithms," CCECE, May 2005, Saskatoon, SK, Canada.

In previous chapters, some background information on the applications, technologies, and system model for passive acoustic source localization were introduced. As a starting point, this chapter focuses on the selection of a preferred localization algorithm for the application of interest. There have been numerous methods proposed during the past century. Here only the most representative and relevant algorithms are considered. The system model defining the source, sensors, signal, noise, and wave propagation adopted for all three algorithms compared in this chapter is the same as that given in Chapter 3, though different data processing techniques are used. A brief description is provided for each algorithm, and then performance evaluation and comparison are done before the final conclusion is presented.

The manuscript included here has been modified to follow the standard and naming conventions of this thesis. Some wording changes have been incorporated to present the candidate algorithms proposed in the literature with better clarity. Some extra simulation done after the original paper was published has also been added in this chapter with the results being rendered in Fig. 4.5. All results presented in the original manuscript remain valid.

Performance Comparison for Three Source Localization Algorithms

Sha Li*
sha.li@usask.ca
University of Saskatchewan

Brian L.F. Daku,
brian.daku@usask.ca
University of Saskatchewan

Abstract

This paper selects and compares three representative source localization algorithms discussed in the literature to identify the preferred one for the application of microseismic event localization. One candidate is a classical two-step location estimator referred to as Cross-Correlation-Simplex in this paper. Another candidate is a one-step algorithm referred to as Energy-based Grid Search (EGS) here. The other is also a one-step location estimator named Approximated Maximum Likelihood (AML) proposed in [1]. Monte Carlo simulations are done, and the results indicate that the EGS algorithm is the most robust choice for the application of interest.

Keywords — localization; short-time duration; grid search; energy-based.

4.1 Introduction

Passive source localization that uses sensor array signal processing has been investigated for a variety of applications, such as radar, sonar, navigation, geophysics and acoustic tracking. Here the application of interest involves localizing microseismic events caused by rock rupture and shifting in underground potash mines [2]. Since the possibility of multiple microseismic events occurring at the same time is very low, only the case of single source localization is considered.

For a far-field radiating source, only the Direction of Arrival (DOA) is observable [3, 4]. However, in a near-field scenario, where the sensors are located in the vicinity of the radiating source, both range and bearing direction can be estimated. A classical process for locating a near-field acoustic source involves two steps: first estimate the Time Difference Of Arrivals (TDOAs) between all sensor pairs, then locate the source with the aid of a set of non-linear equations defined by these TDOAs [5,6]. This traditional procedure

does not render satisfying location estimates for signals of short-time duration such as the acoustic signals generated by microseismic events, which have been observed to last roughly about 0.2 second only. In this case, obtaining accurate TDOA estimates can be too challenging a task due to the noise present, which may amplify the error within the TDOA estimates during the following source location estimation procedure.

An alternate approach to the two-step localization mentioned above is discussed in [2], which avoids the separate time delay estimation procedure but directly searches over a grid of hypothesized source locations to find the grid point that best explains the observed measurements. The particular approach described in [2] is given by a function that calculates the total energy from the sum of the sensor signals, each of which is time-shifted by an amount consistent with the hypothesized location of the event. It results in a very robust one-step algorithm that works well with short-time duration signals, which in this paper is referred to as the Energy-based Grid Search (EGS) algorithm. It can be implemented in either hardware or software to yield real-time estimating results, given the recent technology advances in related areas.

In a recently published paper [1], another one-step estimation algorithm, called the Approximated Maximum Likelihood (AML) algorithm is proposed. It transforms the data observed in the time domain into the frequency domain, then based on a traditional maximum likelihood approach for DOA estimation, results in a function that is explicitly coincident with source locations and implicitly coincident with the source signals. This algorithm also requires a grid search to find the point that gives the maximum of a function, a point which corresponds to the estimate of the source location.

In this paper, Monte Carlo simulations are done to compare the performance of three candidate algorithms for the application where a single source in the near field producing a short-time duration signal, is to be localized. The two single-step algorithms are then further compared with each other to provide a close-up look at their performance using a different simulation setup.

It should be noted that the sensor positions and the propagation speed of the signals are both assumed to be known. In the next section, the three algorithms to be compared are briefly introduced. Then the simulation model and procedure is described, followed in the last section giving some preliminary performance results and analysis.

4.2 Algorithm Description

The three candidate algorithms chosen for study in this paper are described in this section.

4.2.1 Cross-Correlation-Simplex Algorithm

As mentioned earlier, a traditional two-step localization algorithm first estimates time delays and then, based on these time delays estimates the source location. The Cross-Correlation-Simplex, chosen as a

representative two-step localization algorithm, first requires partitioning of the sensors into pairs, since it does not estimate the specific time delay of arrival at each sensor, but only the differences between the time delays, referred to as TDOAs.

The technique adopted to estimate the TDOAs by the Cross-Correlation-Simplex algorithm is a classical cross correlation technique. The cross correlation between any two sensor signals, denoted $r_i(t)$ and $r_j(t)$, can be expressed in the form of the crosspower spectral density function with the aid of widely-known Fourier transform:

$$R_{ij}(\delta\tau) = \int_{-\infty}^{\infty} G_{r_i r_j}(f) e^{j2\pi f \delta\tau} df, \tag{4.1}$$

where $G_{r_i r_j}(f)$ is the crosspower-spectrum of $r_i(t)$ and $r_j(t)$. The time argument $\delta\tau$ at which the correlator achieves a maximum is the TDOA estimate.

In practice, the two sensor signals are usually random processes since they contain additive random noise. Hence only an estimate $\hat{G}_{r_i r_j}(f)$ of $G_{r_i r_j}(f)$ can be obtained from finite observations of $r_i(t)$ and $r_j(t)$ and the integral

$$\hat{R}_{ij}(\delta\tau) = \int_{-\infty}^{\infty} \hat{G}_{r_i r_j}(f) e^{j2\pi f \delta\tau} df \tag{4.2}$$

is evaluated and used for estimating the TDOA.

In the case of multiple sensors, the above expression is used for each pair of sensors, and a group of TDOA estimates results. As pointed out in [7], there is art in how the sensors can be partitioned. Choosing a single sensor to appear in all sensor pairs puts excessive emphasis on that sensor. This is equivalent to assigning an 100% observation certainty over the signal received at the reference sensor. If for some reason there is a large error in the signal arrival time recorded at the reference sensor, the estimate of the source location can be even more erroneous. To eliminate this problem, the partitioning can be done such that no single sensor appears in more than one sensor pair. However, there is a shortcoming of this strategy. It limits the number of TDOAs that can be used later for source localization. Another strategy, as proposed in [7], is to use the mean of the estimated arrival times as the reference value to be subtracted from the arrival time at each sensor. Normalizing the estimated arrival times to their respective mean results in a set of values that bears more statistical consistency.

With the group of TDOAs obtained as described above, the next step is to estimate the source location. The method selected here to calculate this estimate is called the simplex method, originally proposed by Nelder and Mead in 1965 [8]. Its basic idea is to use a general simplex that can move and contort to give a best-fit solution that minimizes a function with K variables in a given error space. For the application of interest, the K variables are the K TDOAs calculated, the function to minimize is the misfit between the calculated and estimated TDOAs, and the error space is made up of all the possible values of the misfit.

How the misfit should be calculated is critical since it affects the accuracy of the final location estimation.

Generally speaking, there are two expressions to use to calculate the error. One is the median value statistic, or $L1$ norm, whose expression is given by

$$Err_{L1} = \frac{1}{K}\sum_{k=1}^{K} |\delta\tau_{est_k} - \delta\tau_{cal_k}| \tag{4.3}$$

when Err_{L1} is a minimum, and where $\delta\tau_{est_k}$ and $\delta\tau_{cal_k}$ are the estimated and calculated TDOA measurements, respectively, K is the number of TDOAs to compare and "$|\ \ |$" denotes absolute value. The other expression often used is the least squares, or $L2$ norm, given by

$$Err_{L2} = \frac{1}{K}\sqrt{\sum_{k=1}^{K} |\delta\tau_{est_k} - \delta\tau_{cal_k}|^2} \tag{4.4}$$

when Err_{L2} is a minimum. However, $L2$ norm weights each TDOA misfit by the square of its error, hence a single value with a large error has the potential to introduce a big bias into the location calculation, resulting in poor location estimates, especially for a limited number of TDOA observations. For example, picking up noise spikes or confusing S and P wave arrivals at one sensor can lead to such a problem. Hence $L2$ norm is not a good choice. $L1$ norm, in comparison, does not suffer as much a performance degradation in the minimization procedure even when there are more than one larger errors existing in the TDOA estimates, so long as the number of good TDOA estimates is greater. This makes it a more robust way to calculate the source position and is therefore chosen as the error function to be minimized in this paper.

The simplex algorithm has been frequently used to solve all kinds of fitting-curves-to-data problems since computers were adopted for numerical calculation tasks. A geometric simplex is a figure having one more vertex than the number of dimensions in which it is defined. A simplex figure is a triangle when defined within a two-dimensional space and a tetrahedron within a three-dimensional space, etc. The vertices of a simplex figure represent values in an error space. The numerical searching procedure with the aid of a simplex for microseismic localization application can be implemented as follows:

1. Pick at least two estimated TDOAs from the results obtained using the cross-correlation technique, e.g., from the sensors with the highest SNR values.

2. Make an initial guess of the source location at a minimum of three points. Initial position choices depend on the distance of the event from the geophone array and the dimensions of the array.

3. For each of the three or more test points, obtain the set of calculated TDOAs according to the same sensor partition used for the TDOA estimation step. Then normalize the set of calculated TDOAs to their mean.

4. Calculate the Err_{L1} for each test point by comparing the calculated TDOAs to the estimated TDOAs.

5. Determine which of these vertices has the highest value for Err_{L1} and which has the lowest value. Replace the high Err_{L1} point with another lower error location position. New points for the simplex are

found by testing "reflected," "expanded," "contracted," and "shrunken" vertices, which are depicted for the two-dimensional (three-point simplex) case in Figure 4.1.

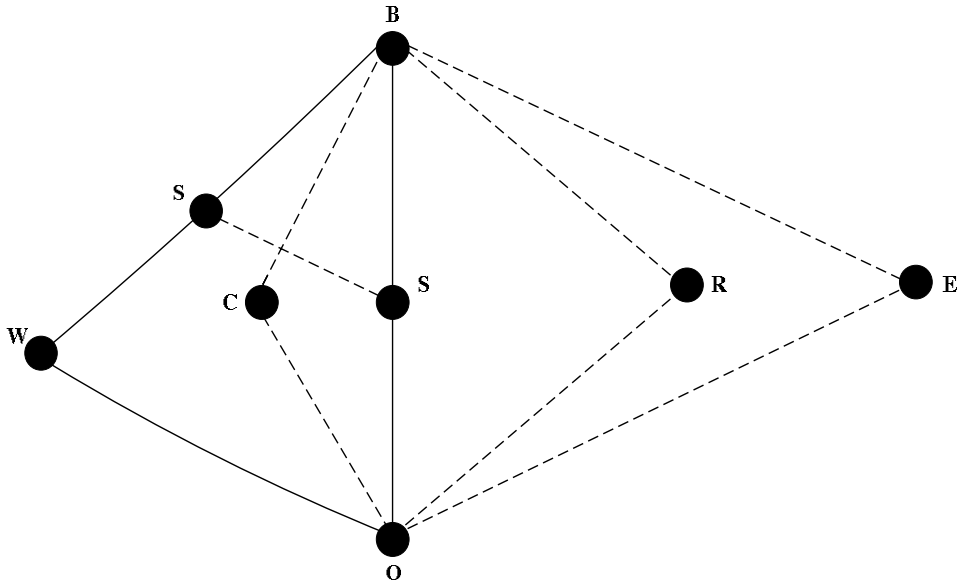

Figure 4.1: Three-point two-dimensional simplex. Four mechanisms of movements are indicated: reflection, expansion, contraction, shrinkage. B = best vertex, W = worst vertex, R = reflected vertex, E = expanded vertex, C = contracted vertex, and S = shrunken vertex.

The pattern of test points is thus moved away from points with larger errors, toward points with smaller errors.

6. Continue moving the simplex test points until the location point gives the best-fit, or least error, when compared with the measured data. This is the location where theoretical/calculated TDOAs are very close to the true TDOAs estimated: the minimum in the error space which represents the best-fit microseismic event location.

In practice, simplex iteration stops at a point where Err_{L1} meets a set of minimum value conditions. The conditions may be that Err_{L1} is less than a certain pre-established value, that the step size in the search has fallen below a certain value, or that a certain maximum number of trial locations has been reached. Though establishing a minimum residual error threshold beforehand is a common technique used to stop a numerical search, the iteration is stopped in the simulation here when the simplex figure has contracted such that all vertices are physically close enough to one another according to a certain closeness criterion.

4.2.2 Approximated Maximum Likelihood Algorithm

The second candidate algorithm to evaluate is proposed in [1] as an optimal parametric approach to solve the problems of wideband source localization in the near field. This algorithm, called Approximated Maximum Likelihood (AML) algorithm in [1], is developed based on the classical maximum-likelihood DOA estimator for narrowband signals. The discrete signal is first mapped into the frequency domain, where its wideband spectrum is constructed using the narrowband model for each frequency segment predefined. The AML algorithm is applied to each positive frequency segment to estimate the source location and the AML metric produces a combination of all the frequency segments that yields the final location estimate.

A block of L discrete samples in each sensor signal can be mapped into the frequency domain with a DFT of length N. It is known that DFT does circular time shift, which creates error in the actual time shift when L is small. Hence the impact of the L value being small on the performance of the AML will be evaluated.

The frequency domain M-sensor array signal model can be given by

$$\vec{X}_r(n) = \vec{D}(n)X_s(n) + \vec{X}_n(n), \tag{4.5}$$

for $n = 0, ..., N-1$, where the received signal spectrum is given by $\vec{X}_r(n) = [X_{r_1}(n)\ X_{r_2}(n)\ ...X_{r_M}(n)]^T)$; the propagation vector is given by $\vec{D}(n) = \left[e^{-j2\pi n\tau_1/N}\ ...e^{-j2\pi n\tau_M/N}\right]^T$, τ_m the fractional time-delay in samples, as the authors of [1] assume uniform signal gain level at all the sensors in their actual derivation; the source signal spectrum is given by $X_s(n)$; and the noise spectrum vector $\vec{X}_n(n) = [X_{n_1}(n)\ X_{n_2}(n)...X_{n_M}(n)]^T)$ is zero-mean complex white Gaussian distributed with variance $N\sigma^2$ in each element.

The zero frequency and negative frequencies can all be ignored. Hence by heaping together only the $N/2$ positive frequency subbands, as defined in (4.5) into a column, all the received data can be rewritten into a $NM/2 \times 1$ space-temporal frequency vector as $\vec{X}_r = g(\vec{\Theta}) + \vec{\xi}$, where $\vec{\Theta}$ denotes the unknown parameter vector $[\vec{p}_s, X_s]$ for the single source scenario, $g(\vec{\Theta}) = \vec{D}(n)X_s$, and $E[\vec{\xi}\vec{\xi}^H] = N\sigma^2 I_{NM/2}$.

The log-likelihood function of the complex Gaussian noise vector $\vec{\xi}$, after ignoring irrelevant constant terms, is given by $L(\vec{\Theta}) = -\parallel \vec{X}_r - g(\vec{\Theta}) \parallel^2$. The maximum-likelihood estimation of the source locations and source signals is given by the following optimization criterion:

$$\underset{\vec{\Theta}}{max} L(\vec{\Theta}) = \underset{\vec{\Theta}}{min} \sum_{n=1}^{N/2} \parallel \vec{X}_r(n) - \vec{D}(n)X_s(n) \parallel^2, \tag{4.6}$$

which is equivalent to finding $\underset{\vec{\Theta}}{min} f(n)$ for all n bins, where

$$f(n) = \parallel \vec{X}_r(n) - \vec{D}(n)X_s(n) \parallel^2 . \tag{4.7}$$

The minima of $f(n)$ with respect to the source signal vector $X_s(n)$ must satisfy $\frac{\partial f(n)}{\partial X_s(n)} = 0$; hence after

some further arrangement, the AML source location estimate in the single-source case becomes $J(\vec{p}_s) = \sum_{n=1}^{N/2} |\mathbf{B}(n, \vec{p}_s)|^2$, where $\mathbf{B}(n, \vec{p}_s)$ is the beam-steered beamformer output in the frequency domain [9].

[5] shows that the AML approach is identical to the maximum weighted cross-correlation approach for the single-source and near-field case when using frequency spectrum representation, and that the maximum weighted cross-correlation criterion, denoted J_{wcc}, maximizes the sum of the weighted cross-correlation functions with the relative time delays for different parameters $\vec{p}_s$, the source positions.

$$\begin{aligned} & J_{wcc}(\vec{p}_s) \\ = & \frac{1}{N} \sum_{n=1}^{N-1} \sum_{m=1}^{M} X_{rm}(n) e^{-j2\pi t_m(\vec{p}_s)\frac{n}{N}} \sum_{l=1}^{M} X_{rl}(n) e^{-j2\pi t_l(\vec{p}_s)\frac{n}{N}}, \\ = & \frac{1}{N} \sum_{n=1}^{N-1} |\mathbf{B}(n, \vec{p}_s)|^2. \end{aligned} \tag{4.8}$$

4.2.3 Energy-based Grid Search Algorithm

The third candidate to evaluate is the so-called Energy-based Grid Search algorithm that has been studied in [2]. The microseismic event is assumed to take place within a finite three-dimensional space to be investigated. The algorithm performs a global search over a grid defined within such a finite space, which can be abstracted as a three-dimensional box. The grid points represent hypothesized source positions. An energy function is then computed for each grid point inside the box, and the point that maximizes the energy function yields the estimate of the actual source position. The energy function is given by

$$W(\vec{\tilde{p}}_s) = \int_0^T \left[\sum_{m=1}^{M} r_m(t + \tilde{\tau}_m) \right]^2, \tag{4.9}$$

where $\vec{p}_s$ represents the real source position, $\vec{\tilde{p}}_s$ the hypothesized source position, and $\tilde{\tau}_m$, given by $\tilde{\tau}_m = \frac{\|\vec{\tilde{p}}_s - \vec{p}_m\|}{c}$, the time an acoustic wave takes to reach sensor m from its assumed origin, the hypothesized position. The integration limits in practice can be obtained from the strongest signal received at the sensor closest to the event, hence the assumption of the integration limits having negligible estimation error is reasonable.

(4.9) indicates that the received sensor signals, $r_m(t)$, m = 1,..., M, are shifted back in time by $\tilde{\tau}_m$, m = 1,..., M, according to the distance between the grid point $\vec{\tilde{p}}_s$ and sensor position $\vec{p}_m$. The shifted sensor signals are then summed, squared, and integrated to give the final output value. The grid point $\vec{\tilde{p}}_s$ at which (4.9) achieves its maximal value represents the estimate of the true source position $\vec{p}_s$. (4.9) is a function of the grid point, but the shifted signal is not scaled in its amplitude accordingly. The mathematical foundation of this algorithm lies in the properties of signal auto-correlation. Though in fact (4.9) is cross-correlation between sensor signals, it can be interpreted as auto-correlation between two copies of the source signal though with different amplitudes. When there is no noise, cross-correlation peaks when the time lag

between two signals is zero. For (4.9), this happens only where the hypothesized source position falls onto the true one, when each sensor signal is shifted a correct amount of time back to the origin.

This total energy approach given by (4.9) is conceptually easy and does not distort any original data during calculation but makes direct use of them. This feature makes it much more robust and reliable, compared with other algorithms that require various types of preprocessing of the data.

4.3 Simulations and Performance Evaluation

In the simulation model, the passive sensors are assumed to be omnidirectional, having identical responses and the same i.i.d. white Gaussian noise. Compared with the original source signal, sensor signals experience time delays, as well as amplitude degradation that is inverse to the distance, and the additive noise.

All the simulations are done using Matlab. A signal with random phases but fixed-frequency components and energy is generated to serve as the source signal. The number of data samples is 200, and the sampling frequency is set at 1 kHz, which makes the duration of the signal 0.2 second. The frequency components are set to range from 5 to 100 Hz, reflecting common components of actual microseismic source signals in the application of interest. The speed of propagation used is 345 m/s. At the arbitrarily placed sensors, all signals received are simulated with appropriate time delays and amplitude attenuation. For the EGS and the Cross-Correlation-Simplex algorithms, conventional time delays are applied, while circular-shift delays are introduced for the AML algorithm.

The first simulation scenario considers five arbitrarily placed sensor and various source locations along a line. This source-sensor geometry is depicted in Figure 4.2.

Appropriate values of SNR are set for all the sensors for each source location along the line in Figure 4.2, based on the assumption that the SNR at the closest sensor always equals 0 dB. For each source position indicated in Figure 4.2, 1000 iterations of simulation are run to obtain the location estimate. The root mean square(RMS) of the estimation error is then calculated and plotted in Figure 4.3.

Results presented in Figure 4.3 show that the two-step Cross-Correlation-Simplex algorithm yields the worst performance among the three. For the Cross-Correlation-Simplex method, the accuracy of the solution depends on two factors. The first is the closeness between the true source position and the initial guessed source positions to start the iteration. The simplex can be trapped in a local minimum if the initial points are too far away from the real source point, or there are simply too many local minimums existing around the true source position. The second factor is the closeness between the observed data (estimated TDOA values) and the real data (true TDOA values), since the solution converges at the point reflected only by the observed data. The closer the observed data are to the real data, the more accurate the final solution will be. The simplex method represents a non-linear relationship between the TDOAs and the source position.

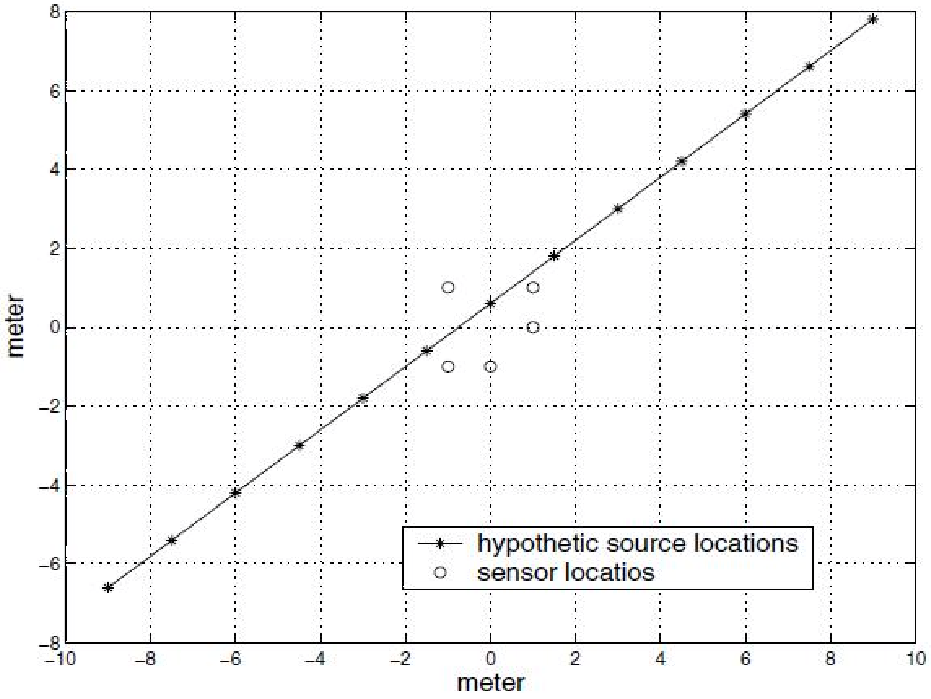

Figure 4.2: Sensor configuration and assumed source event locations for simulation 1

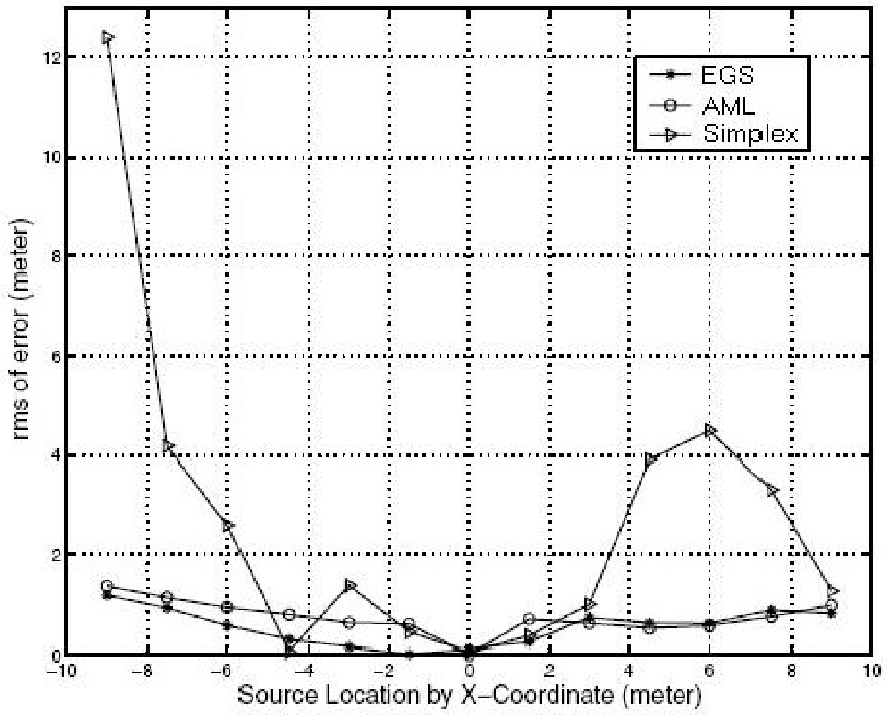

Figure 4.3: Comparison between the three candidate localization algorithms

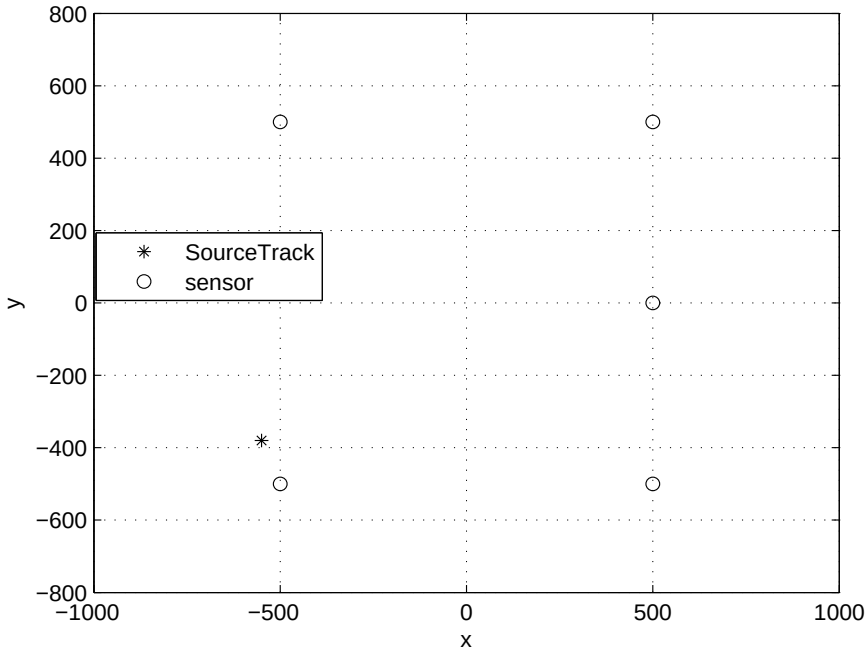

Figure 4.4: Sensor configuration and assumed source event locations for simulation 2

When only 200 data samples from a source signal lasting only 0.2 second are available, as in the simulation scenario here, or when the SNR is quite low, poorly estimated TDOA values may well be seen. Hence even if a good guess of the true source location for the initial iteration can be made in such cases, an even poorer source location estimate is still likely to result since the errors in the TDOA estimates can be enlarged during the non-linear procedures of the simplex method. The simulation results confirmed this, and the need for alternative methods to estimate the source location consistently with more accuracy is also validated.

As for the other two single-step algorithms, the EGS algorithm demonstrates an overall better performance compared with the AML in Figure 4.3.

As discovered in [1], the AML algorithm introduces a few artifacts, such as the edge-effect problem for the actual linear time shift due to the inherited circular-shift property of the DFT, and the loss of the orthogonality of the DFT transform when zero padding is done to the original signals. In the latter case, the white Gaussian noise shows correlation across the frequency domain. The situation tends to get worse when the SNR is low. In comparison, the EGS algorithm does not have these problems since it only deals with the original time-domain data. So, theoretically, the EGS algorithm is expected to demonstrate a more robust performance. To verify this, a second simulation is run solely to evaluate the two one-step algorithms for a series of SNR values, assuming a specific source position. The source sensor geometry for the second simulation scenario is depicted in Figure 4.4. For each SNR value, 1000 iterations of simulation are done

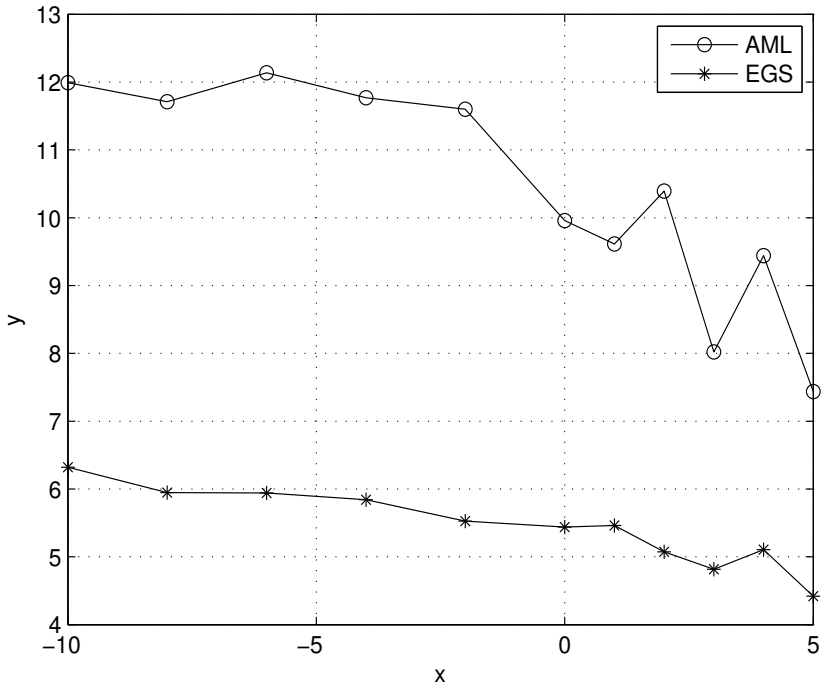

Figure 4.5: Comparison between the EGS and AML

to obtain the location estimate, and the RMS of the estimate error is then calculated to give the results as plotted in Figure 4.5.

Figure 4.5 shows that the EGS algorithm does outperform the AML algorithm. This is primarily because that the number of data adopted in the simulation is only 200, which is not enough for the AML to do DFT with only negligible numeric errors. The EGS algorithm, on the contrary, does not require any preprocessing of the data, which makes it a more robust source localization algorithm even under low SNR values and with a small number of data available. The two algorithms are both unbiased and are supposed to converge at high SNR values without a significant difference in their performances, so simulations are not performed for SNR values higher than 5 dB.

4.4 Conclusion

In this chapter, the one-step EGS localization algorithm is compared with the two-step Cross-Correlation-Simplex algorithm and the one-step AML algorithm. The simulation results show that all three algorithms are capable of estimating the source position, though with varying accuracy. The performances of all three algorithms degrade when the source gets further away from the sensor array. This indicates that the more symmetrically the sensors are arranged around the source, the more accurate the source location tends to

be, given the same SNR value. It is seen that for short-time duration signals, the two one-step algorithms do a better job in estimating the source location than the two-step Cross-correlation-Simplex algorithm does.

The AML algorithm requires DFT processing of the received data, thus it risks larger computational errors when the number of data samples is limited, while the two-step algorithm risks introducing extra TDOA estimation errors that have a non-linear relationship with the source position. Though the AML algorithm is a good one-step algorithm when the signals last longer or the SNR is higher, the EGS demonstrates an overall more stable and robust performance than the AML, as indicated by the simulation results obtained here. Also, the EGS algorithm is conceptually easy and straightforward to implement in both software and hardware. Therefore, the EGS is selected as the preferred algorithm for the application of interest.

Acknowledgments

Sha Li would like to thank Prof. Daku for all his guidance and suggestions during her research work. She would also like to express her appreciation to TRlabs Saskatoon for the funding, the wonderful working environment and the nice and helpful people there.

References

[1] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. on Signal Processing*, vol. 50, pp. 1843–1854, August 2002.

[2] B. L. Daku, J. E. Salt, and S. Li, "An algorithm for locating microseismic events," in *Proceedings of CCECE04*, May 2004.

[3] R.O.Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagation*, vol. AP-34, pp. 276–280, 1986.

[4] I. Ziskind and M. Wax, "Maximum likelihood localization of multiple sources by alternating projection," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1553–1560, October 1988.

[5] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. on Signal Processing*, vol. 42, pp. 1905–1915, August 1994.

[6] M. Brandstein, J. Adcock, and H. Silverman, "A closed-form location estimator for use with room environment microphone arrays," *IEEE Trans. Speech, and Audio Processing*, vol. 5, no. 1, pp. 45–50, 1997.

[7] A. F. Prugger and D. J. Gendzill, "Microearthquake location: a nonlinear approach that makes use of a simplex stepping procedure," *Bulletin of th Seismological Society of America*, vol. 78, pp. 799–815, April 1988.

[8] J. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.

[9] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

5 Cramér-Rao Bound on Passive Source Localization for General Gaussian Noise

The content was originally published as::

S. Li and B. L. Daku, Cramér-rao bound on passive source localization for general gaussian noise, IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences, vol. E93-A, pp.914-925, May 2010.

The previous chapter presents the preferred localization algorithm for the application of interest after comparing three representative candidate algorithms given in the literature. The EGS is chosen as preferred due to its robustness and therefore becomes the foundation of the work carried in chapters 5 - 7 to follow. This chapter focuses on the development of closed-form CRB expressions for stationary Gaussian noise cases, though the mathematical procedure can be applied to non-stationary Gaussian noise cases as well.

Cramér-Rao Bound on Passive Source Localization for General Gaussian Noise

Sha Li* and Brian L.F. Daku

Abstract

This paper focuses on the development of Cramér-Rao Bound (CRB) expressions for passive source location estimation in various Gaussian noise environments. The scenarios considered involve an unknown deterministic source signal with a short time duration, and additive general Gaussian noise. The mathematical derivation procedure presented is applicable to non-stationary Gaussian noise problems. Specifically, explicit closed-form CRB expressions are presented using the spectrum representation of the signal and noise for stationary Gaussian noise cases.

Keywords — CRB, source localization, estimation error variance

5.1 Introduction

There has been great interest in source localization using passive sensor arrays in various applications, such as underwater acoustic applications [1], video conferencing [2, 3], and geophysics applications [4, 5]. A variety of algorithms have been developed to cater to these applications. The corresponding Cramér-Rao Bounds (CRBs), widely known as the theoretical lower performance bounds for estimation algorithms, have also been developed for these applications. For source localization in far-field cases, only the direction of arrival (DOA) can be estimated. For near-field cases, exact source locations can be determined using traditional two-step localization techniques that rely heavily on time difference of arrival (TDOA) [6, 7]. In the cases of using TDOA estimates, the least square and hyperbolic intersections techniques [8–11] are the most popular methods for obtaining the final source location estimate. Hence, it is not surprising that many CRB expressions developed for source localization problems in the literature, such as [12–14], are in terms of TDOA/DOA estimation errors.

Both authors are with the University of Saskatchewan, Department of Electrical Engineering, 57 Campus Drive, Saskatoon, SK, Canada, S7N 5A9. Emails: sha.li@usask.ca brian.daku@usask.ca.

In all of the applications mentioned above, the source signals being localized are assumed to have a time duration long enough, which is a condition necessary for those algorithms to work above threshold. In contrast, the object of study here is signals with a short-time duration and finite energy. The specific application of interest is the positioning of local, low-intensity earthquakes that are referred to as microseismic events in the surrounding area of an underground mine [15]. Hence, it is categorized as a near-field problem. Though similar to other geophysics applications involving ordinary seismic events, a different signal model is required, since microseismic events usually last much less than a second. It has been found in [15] that the one-step estimator avoiding the TDOA estimation procedure is more robust than the two-step estimators relying heavily on TDOA estimates, when the time-bandwidth product of the source signal is fairly small. Naturally, a CRB expression needs to be developed to evaluate this class of one-step, direct algorithms for localization of the source signal concerned. Since the mining rooms of interest are straight tunnels that expand in an arbitrary direction, this paper focuses on the source location estimation in a given direction.

There have been explicit CRB expressions presented for various parameter estimation problems besides TDOA/DOA. For example, CRB expressions for estimation of channel and signal parameters or of the signal itself are considered for additive white Gaussian noise (WGN) in [16–18]. In [19] and [20], CRB expressions are presented for estimating discrete frequency components in the presence of special coloured Gaussian noise (CGN) that can be represented linearly by functions of WGN. Though providing some insight into the CRB development procedure, the above references are not directly related to the application of interest in this paper. [21] gives a CRB expression for the final location estimation error that falls in this paper's area of interest, but it is derived using the assumption that the error variance introduced by previous TDOA estimates is a given. This makes it applicable only to cases where long signal records are available such that the TDOA values can be estimated with acceptable error. [22] focuses on the source localization application using a single-step source location estimator based on maximum likelihood techniques. However, it treats the problem in such a way that the sensor signal amplitude is independent of the distance between the source and the sensor. This is not appropriate for most near-field applications. Though [22] tates it considers the case where the source position and source signal samples are both unknown, the contribution of unknown source signal samples to the CRB for the distance error between the estimated location and the true source location is actually ignored during the derivation. In addition, [22] only considers WGN and discrete time signals.

One important feature of the source localization problem lies in the characterization of the noise assumed for each observed sensor signal. If there is no a priori information available about the noise or if the noise is known to be WGN, it is sufficient to adopt the white Gaussian model. But in many cases, a stationary CGN model may better reflect the physical situation. For example, in cases where there exist electronic elements, such as antennas or radio frequency/microwave filters, which shape the original noise spectrum before the noise enters the data-processing part, and also, in cases where there is an interfering signal that

can be characterized as a Gaussian process. When these unwanted random processes can be estimated and characterized statistically or, in the best situation, when they are simply already known, proper filters should be developed accordingly to reduce such noise in order to improve the estimation/detection accuracy. As a consequence, a corresponding CRB expression to evaluate an estimator applied to the stationary CGN case is highly desirable.

For WGN, discrete-time samples can be utilized directly to construct the probability density function (p.d.f.) of the observations and hence the log-likelihood function can be easily generated. The independence between samples greatly simplifies the CRB expression derivation. However, for continuous sensor signals that contain coloured noise components, the sampling procedure will not help simplify the procedure leading to the CRB expression, since the samples are not independent of one another. Obviously, some means are needed to ease the difficulty raised by the more complex characteristics of CGN. Here the Karhunen-Loève (K-L) expansion is the technique chosen.

The well known Fourier transform (FT) resembles the K-L expansion. Its fundamental purpose is to decompose a signal into the sum of weighted sinusoidal functions that extend from negative infinity to positive infinity in time [23] Despite its simple interpretation of pure frequencies, the FT is not always the best tool to analyze the majority of real life signals. The classical linear FT and the square of the FT do not reveal the important characteristics of the signals whose frequency component alters over time. Common examples for such signals that are usually of finite or even relatively shorter duration include biomedical signals [24–27] and seismic signals [28]. Microseismic signals of interest here, usually last a fraction of one second. For estimation problems involving such signals, the sinusoidal functions are not good models. Hence a mathematical technique other than the FT is needed to address this issue. Such a technique shall also be useful in the derivation of a corresponding theoretical lower performance bound for the purpose of localization algorithm evaluation. To meet this goal, the major mathematical approach presented here is applicable to CRB development for transient source signals in the presence of non-stationary Gaussian noise. However, deriving a set of orthonormal basis functions to expand an arbitrary non-stationary Gaussian process in order to obtain an analytical CRB is beyond the scope of this paper.

In this work, the sensor noise is initially assumed spatially correlated and the corresponding CRB development procedure is presented first. Since sparse sensor placement is employed in the specific application of microseismic event localization, spatially uncorrelated sensor noise is then considered, demonstrating more compact analytical CRB expressions, as expected. These closed-form CRB expressions are presented as functions of the spectrum representation of the signal and noise for three cases: stationary CGN with a different spectrum envelope at each sensor, stationary CGN with the same spectrum envelope at each sensor, and WGN at each sensor.

In [29], the CRB expression is derived using Fourier-transformed observations in the frequency domain

from the beginning. This implies that the result is applicable to the spatially uncorrelated coloured stationary Gaussian noise case, a case considered in this paper. However, the derivation procedure adopted in [29] is based strictly on the FT, which cannot be generalized to treat non-stationary Gaussian noise problems. In contrast, the mathematical technique presented in this paper, which is based on the Karhunen-Loève (K-L) expansion, can be used to address such problems, since Mercer's theorem ensures the existence of a certain orthonormal function set that can expand a positive definite kernel bilinearly with uniform convergence. Analytical orthonormal functions have already been solved in the literature [30] and [23] for some special forms of non-stationary processes, like the Wiener-Levy and the Brown-Bridge processes, though solving analytical orthonormal functions is not always possible. Also, [29] treats the source signal as a Gaussian random process and focuses on far-field applications, which are quite different from the application of interest here. Furthermore, it assumes that the source signal and noise are both known, which simplifies the mathematical derivation but cannot satisfactorily reflect the application scenario considered here. In the work here, closed-form CRB expressions will only be developed for the case of stationary Gaussian processes using the adopted mathematical technique. CRB expressions for the case of non-stationary Gaussian processes will be discussed in future work to limit the length of this writing.

The paper is arranged in the following manner. A basic system model description concerning the noise, signal, signal propagation, and sensor configuration is given first. Next, the detailed CRB development for source position estimation in a given direction is presented. This includes second moment characterization of the observations, discussion of the special format of the Fisher information matrix (FIM), and derivation leading to the final closed-form CRB expressions for the aforementioned specialized noise environments. Then after a graph of the mathematical results is presented, the paper ends with a summary and conclusions.

5.2 System Model

The basic physical model adopted in this paper makes similar assumptions as in [15]: the signal source is a single-point; the propagation medium is homogeneous and elastic (lossless) that enables spherical spreading, where the signal energy attenuates proportionally to the square of distance in all directions; all indirect-path signals are weak enough not to be considered; all the sensors adopted respond identically; the output of each sensor is continuous and contains additive Gaussian noise that can be spatially correlated; and the noise at each sensor has a zero-mean but a different unknown auto-covariance function.

The spatial coordinate system used here is the three dimensional x, y, z Cartesian system. The source position is generally given by $\vec{p}_s = (x_s, y_s, z_s) = \epsilon_s \vec{e}_s$, where the unit vector $\vec{e}_s$ indicates the source angle and the scalar ϵ_s indicates the source distance.

There are M sensors that are placed arbitrarily and sparsely in the system. The position of sensor m is

denoted $\vec{p}_m = (x_m, y_m, z_m)$. The distance from the source to sensor m is denoted

$$\begin{aligned} d_m &= \| \vec{p}_m - \vec{p}_s \| \\ &= \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2 + (z_m - z_s)^2}. \end{aligned} \tag{5.1}$$

The output of sensor m, denoted $r_m(t)$, is a continuous sample function of the stochastic process $\mathbf{r}_m(t)$, which is given by

$$\mathbf{r}_m(t) = s_m(t) + \mathbf{n}_m(t), \tag{5.2}$$

where $s_m(t)$ is the signal received at sensor m, and $\mathbf{n}_m(t)$ is an additive random noise process. $s_m(t)$ is simply a delayed and attenuated version of the original unknown deterministic source signal $s_0(t)$, which is independent of $\mathbf{n}_m(t)$. $\mathbf{n}_m(t)$ is Gaussian noise with a zero-mean and unknown covariance. The random noise process at each sensor may be non-stationary and spatially correlated. However, when sensors are sparsely positioned, the noise can be taken as approximately spatially uncorrelated.

It is assumed that all received signals are collected by sensors that are at least one unit distance away from the source to accommodate the signal attenuation model adopted in this paper. The unit can be arbitrary so long as it is in accordance with the distance measure system used for the mathematical derivation procedure. Now define

$$s_m(t) = \frac{s_0(t - \tau_m)}{d_m}, \tag{5.3}$$

where $\tau_m = \frac{d_m}{c}$ is the time for the unknown signal $s_0(t)$ to travel from the source location to sensor m, and c is the propagation speed of the mechanical wave.

It is physically reasonable to assume that the source signal $s_0(t)$ is real, contains no discontinuity, and is finite with a time duration T. The start time of the source event is taken as the time origin, and T_o is the end of the observation time, which is assumed to be much greater than T. To facilitate mathematical development, T_o may be extended to infinity. $n_m(t)$ and $r_m(t)$ are both real functions and denote sample functions of $\mathbf{n}_m(t)$ and $\mathbf{r}_m(t)$ respectively. Expressed in column vectors, the following definitions are given

$$\begin{aligned} \vec{N}(t) &= [n_1(t) \quad n_2(t) \quad \cdots \quad n_M(t)]^T, \\ \vec{S}(t) &= [s_1(t) \quad s_2(t) \quad \cdots \quad s_M(t)]^T, \\ \vec{R}(t) &= [r_1(t) \quad r_2(t) \quad \cdots \quad r_M(t)]^T. \end{aligned}$$

With "$E[\ ]$" representing the expected value, the noise mean vector denoted $\vec{M}(t)$, and the matrix of noise cross covariance functions denoted $\mathbf{C}_{NS}(t_1, t_2)$, are given below respectively,

$$\begin{aligned} &\vec{M}(t) = E[\vec{N}(t)] = 0, \\ &\mathbf{C}_{NS}(t_1, t_2) = E\left[\vec{N}(t_1)\vec{N}(t_2)^T\right]. \end{aligned}$$

The covariance between $n_m(t)$ and $n_n(t)$ at times t_1 and t_2 respectively makes an element of the matrix $\mathbf{C}_{NS}(t_1, t_2)$, and is referred to as $\mathbf{C}_{NS}^{mn}(t_1, t_2)$.

In the rest of the study, superscript "$*$" is used to represent the complex conjugate, superscript "T" and "H" to represent the simple transpose and conjugate transpose respectively, "diag[ ]" to represent a diagonal matrix, and "./", ".$\times$" and ".^{2}" to represent element-wise division, element-wise multiplication and element-wise square, respectively.

5.3 Development of the Cramér-Rao Bound

In general source localization applications, the parameter to be estimated is the coordinate of the source position. Hence the CRB in correspondence is a vector. In the application of microseismic event localization for mining safety, the long and straight mine rooms or tunnels involved may expand in any direction, along which the evaluation of the source location estimation error is critical. In this scenario, the desired CRB becomes just a scalar that reflects how close the estimated and true source positions can be in the given direction.

Denote $\hat{\vec{p}}_s$ the estimated $\vec{p}_s$, $\vec{e}$ the arbitrary mine room direction, and ϵ the estimation error along direction $\vec{e}$, to give $\epsilon = \| \vec{p}_s - \hat{\vec{p}}_s \|$. Since only ϵ, not ϵ_s or $\vec{e}_s$, is of interest, the origin of the coordinator can be moved to the true source location to give $\vec{p}_s = (0, 0, 0)$, and hence

$$\epsilon = \| \hat{\vec{p}}_s \| = \sqrt{\hat{x}_s^2 + \hat{y}_s^2 + \hat{z}_s^2}, \tag{5.4}$$

$$d_m = \| \vec{p}_m \| = \sqrt{x_m^2 + y_m^2 + z_m^2}. \tag{5.5}$$

In this section, three issues are treated given (5.4) and (5.5). The first is to prepare the observed continuous waveforms, so that the joint p.d.f. can be constructed. This is achieved by mapping a continuous random process from the time domain to a new domain whose coordinates are represented by a set of orthonormal functions. In this new domain, a set of individual random variables is obtained to equivalently represent the original time process. The second issue is to study the structure of the FIM under specified conditions given in this paper in order to avoid unnecessary mathematical derivation. Finally, the detailed CRB development is presented.

In the sequel, without special notification, indices i, k are used to number the axes of the new coordinate system and m, n are used to number the sensors, or equivalently, noise random processes.

5.3.1 Second Moment Characterization of the Observations

Continuous random processes are not easy to handle using classical approaches for CRB development in detection and/or parameter estimation problems. In cases where the noise can be modeled as a WGN

process, sampled data are useful, since they are discrete and independent random variables. However, for a more general Gaussian random process that can be stationary coloured or even non-stationary, its covariance bears no such favourable feature and its discrete-time samples are not as useful as those of a WGN process. A more general continuous-to-discrete transformation approach other than the sampling procedure is needed and the Karhunen-Loève decomposition is found to serve this purpose very well.

The basic idea of the Karhunen-Loève decomposition [23] is to break up a continuous random process given in the time domain into a series of independent discrete random variables in a new domain. This is realized with the aid of orthonormal basis functions that represent the coordinate axes of the new domain. The resulting discrete random variables are projections of the continuous-time random process on all axes of the new coordinate system and they are obviously orthogonal to each other. The second-moment characterization for an arbitrary continuous-time random process can be easily achieved in terms of such uncorrelated random variables in the new domain, a critical step in simplifying the CRB development procedure in general.

Assume that there is a complete set of orthonormal basis functions $\{\phi_k(t)\}$ $(k = 1, 2, \cdots K)$ that spans the two-dimensional complex space and the continuous-time variable $t \in (0, T_o)$. Note that K might be finite or infinite.

For a vector of random noise sample functions, $\vec{N}(t)$, the following pair of expressions exist:

$$\vec{N}_k = \int_0^{T_o} \vec{N}(t)\phi_k(t)dt, \tag{5.6}$$

$$\vec{N}(t) = \sum_{k=1}^{K\to\infty} \vec{N}_k\phi_k^*(t), \tag{5.7}$$

where $\vec{N}_k$ is the projection of $\vec{N}(t)$ on the axis represented by $\phi_k(t)$.

The set of $\{\phi_k(t)\}$ $(k = 1, 2, \cdots K)$ has to satisfy the following integral equation

$$\begin{aligned} \Lambda_k\delta_{ik} &= E[\vec{N}_i\vec{N}_k^H] \\ &= \int_0^{T_o} \phi_i(t_1) \int_0^{T_o} \mathbf{C}_{NS}(t_1,t_2)\phi_k^*(t_2)dt_2dt_1, \end{aligned} \tag{5.8}$$

where Λ_k is an $M \times M$ eigenvalue matrix and $\phi_k(t)$ represents a scalar eigenfunction of the integral equation; δ_{ik} is the Kronecker Delta.

It is physically reasonable to assume that the noise has non-zero energy. Hence each element of the noise covariance function, $\mathbf{C}_{NS}(t_1, t_2)$, is positive definite and expandable bilinearly according to standard results from linear integral equation theory. In addition, Mercer's theorem promises absolute and uniform convergence of such series expansion. Hence all the noise covariance functions can be written in the matrix format as

$$\mathbf{C}_{NS}(t_1,t_2) = \sum_{k=1}^{K\to\infty} \Lambda_k\phi_k(t_1)\phi_k^*(t_2). \tag{5.9}$$

For Gaussian noise, the matrix of eigenvalues Λ_k is independent of one another for $k = 1, 2, \cdots K$.

Now the noise can be characterized easily in the new domain formed by $\{\phi_k(t)\}$ $(k = 1, 2, \cdots K)$ with the first and second moments that are necessary and complete for Gaussian random processes. As mentioned in section 5.2, $\mathbf{n}_m(t)$ has a zero-mean and thus the mean of the equivalent discrete representation of the noise $\{\vec{N}_k\}$ $(k = 1, 2, \cdots K)$ is zero.

The source signal $s_0(t)$ is modeled as an unknown deterministic signal with finite energy. This enables it to be expanded by the same set of eigenfunctions that are determined by the integral equation whose kernel is $\mathbf{C}_{NS}(t_1, t_2)$. The vector of sensor-received signals, $\vec{S}(t)$, containing delayed versions of the original $s_0(t)$, can be expanded accordingly. It is defined that

$$\begin{aligned} s_k &= \int_0^{T_o} s_0(t)\phi_k(t)dt, && (5.10)\\ s_0(t) &= \sum_{k=1}^{K\to\infty} s_k\phi_k^*(t), && (5.11)\\ \vec{S}_k &= \int_0^{T_o} \vec{S}(t)\phi_k(t)dt, && (5.12)\\ \vec{S}(t) &= \sum_{k=1}^{K\to\infty} \vec{S}_k\phi_k^*(t). && (5.13) \end{aligned}$$

Now the observations have the following expressions

$$\begin{aligned} \vec{R}_k &= \vec{S}_k + \vec{N}_k \\ &= \int_0^{T_o} \vec{R}(t)\phi_k(t)dt, && (5.14)\\ \vec{R}(t) &= \sum_{k=1}^{K\to\infty} \vec{R}_k\phi_k^*(t). && (5.15) \end{aligned}$$

$\{\vec{S}_k\}$ $(k = 1, 2, \cdots K)$ are deterministic, $\{\vec{N}_k\}$ $(k = 1, 2, \cdots K)$ are Gaussian random variables with zero-means and variances denoted $\{\Lambda_k\delta_{ik}\}$, $(i = k = 1, 2, \cdots K)$. It follows that

$$E[\vec{R}_k] = \vec{S}_k, \quad (5.16)$$

$$E[(\vec{R}_i - E[\vec{R}_i])(\vec{R}_k - E[\vec{R}_k])^H] = \Lambda_k\delta_{ik}, \quad (5.17)$$

.

5.3.2 Matrix Format of the CRB

The distance error ϵ along a given direction $\vec{e}$, deterministic source signal components $\{s_k\}$ $(k = 1, 2, \cdots K)$, and the independent noise covariance matrices $\{\Lambda_k\}$ $(k = 1, 2, \cdots K)$ make up the elements of a

column vector $\vec{\Theta}$, denoting the unknown parameter vector, as defined below

$$\vec{\Theta} = [\vec{\Theta}_s \quad \vec{\Theta}_n]^T, \tag{5.18}$$

$$\vec{\Theta}_s = [\epsilon \quad s_1 \quad s_2 \quad \cdots \quad s_K]^T, \tag{5.19}$$

$$\vec{\Theta}_n = [\Lambda_1 \quad \Lambda_2 \quad \cdots \quad \Lambda_K]^T, \tag{5.20}$$

where $\vec{\Theta}_s$ and $\vec{\Theta}_n$ are the unknown vector parameters related to the signal and noise respectively. Due to the assumption of independence between the signal and the noise, $s_0(t, \vec{\Theta}) = s_0(t, \vec{\Theta}_s)$ and $\vec{N}(t, \vec{\Theta}) = \vec{N}(t, \vec{\Theta}_n)$.

The FIM is defined based on the first and second derivatives of the conditional log-likelihood function with respect to each element of $\vec{\Theta}$. In a scenario where both the signal and the noise are unknown but parameterized independently, as assumed here, the FIM has a block matrix form as

$$\mathbf{J}(\vec{\Theta}) = \begin{pmatrix} \mathbf{J}_s & 0 \\ 0 & \mathbf{J}_n \end{pmatrix},$$

where the matrices $\mathbf{J}_s = \mathbf{J}(\vec{\Theta}_s)$ and $\mathbf{J}_n = \mathbf{J}(\vec{\Theta}_n)$ are the FIM's corresponding to the signal and the noise respectively. Hence

$$\begin{aligned} \mathbf{CRB}(\vec{\Theta}) &= \begin{pmatrix} \mathbf{CRB}_s & 0 \\ 0 & \mathbf{CRB}_n \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{J}_s^{-1} & 0 \\ 0 & \mathbf{J}_n^{-1} \end{pmatrix}. \end{aligned} \tag{5.21}$$

Since just the CRB for source location estimation error in a certain direction is of interest and (5.21) is a diagonal matrix, only the $\mathbf{J}_s^{-1}$ term needs to be considered.

5.3.3 Detailed Derivation of $CRB(\epsilon)$

In this subsection, the detailed procedure of finding the lower bound of the source location estimation error in a given direction, denoted $\mathbf{CRB}(\epsilon)$ or $\mathbf{CRB}_s(\epsilon)$ equivalently, is presented. From (5.21) and the definition of $\vec{\Theta}_s$ given by (5.19), it is known that

$$\mathbf{CRB}(\epsilon) = \mathbf{J}_s^{-1}(1,1), \tag{5.22}$$

which will not be contingent on $\mathbf{J}_n$.

Given (5.17), it is seen that in the new coordinate system defined by $\{\phi_k\}$ (k =1, 2, $\cdots$, K), the joint

p.d.f. of the observations conditioned on unknown $\vec{\Theta}_s$ is given by

$$\begin{aligned} & \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\} \\ = & \prod_{k=1}^{K} \mathrm{p}\{\vec{R}_k|\vec{\Theta}_s\} \\ = & \prod_{k=1}^{K} \frac{1}{\sqrt{2\pi \parallel \Lambda_k \parallel}} \exp\left[-\frac{1}{2}(\vec{R}_k - \vec{S}_k)^H \Lambda_k^{-1}(\vec{R}_k - \vec{S}_k)\right], \end{aligned} \tag{5.23}$$

where $\mathbf{R} = [\vec{R}_1 \ \vec{R}_2 \ \cdots \ \vec{R}_K]$ is an $M \times K$ matrix.

Taking the logarithm of the above equation gives

$$\begin{aligned} & \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\} \\ = & \sum_{k=1}^{K} \left[\ln \frac{1}{\sqrt{2\pi \parallel \Lambda_k \parallel}} - \frac{1}{2}(\vec{R}_k - \vec{S}_k)^H \Lambda_k^{-1}(\vec{R}_k - \vec{S}_k)\right]. \end{aligned} \tag{5.24}$$

This log-likelihood function is a random variable in the form of the natural logarithm of the conditional density function of the transformed observations.

It is well known that $\mathbf{J}_s$ is Hermitian, since

$$\mathbf{J}_s = E\left[\frac{\partial \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial \vec{\Theta}_s} \cdot \frac{\partial \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial \vec{\Theta}_s^H}\right]. \tag{5.25}$$

For easier manipulation, another equivalent form of the FIM is used:

$$\mathbf{J}_s = -E\left[\frac{\partial^2 \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial \vec{\Theta}_s \partial \vec{\Theta}_s^H}\right]. \tag{5.26}$$

Before $\mathbf{CRB}(\epsilon)$ can be calculated according to (5.22), each element of $\mathbf{J}_s$ has to be obtained to find its inverse, $\mathbf{J}_s^{-1}$. First observe elements $\mathbf{J}_s(i+1, k+1)$, $1 \le i, k \le K$. It is seen that

$$\begin{aligned} & \mathbf{J}_s(i+1, k+1) \\ = & -E\left[\frac{\partial^2 \ \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial \vec{\Theta}_s(i+1) \partial \vec{\Theta}_s(k+1)}\right] = -E\left[\frac{\partial^2 \ \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial s_i \partial s_k}\right] \\ = & \begin{cases} -E\left[\frac{\partial^2 \ \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial s_k^2}\right], & i = k; \\ 0, & i \ne k. \end{cases} \end{aligned} \tag{5.27}$$

This reduces the $\mathbf{J}_s$ matrix into the following form

$$\mathbf{J}_s = \begin{pmatrix} \mathbf{J}_s(1,1) & \mathbf{J}_s(1,2) & \dots & \mathbf{J}_s(1,K+1) \\ \mathbf{J}_s(2,1) & \mathbf{J}_s(2,2) & 0 \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{J}_s(K+1,1) & 0 & \dots 0 & \mathbf{J}_s(K+1,K+1) \end{pmatrix}$$

which can be easily partitioned into a 2×2 block matrix

$$\mathbf{J}_s = \begin{pmatrix} \mathbf{J}_s^{11} & \mathbf{J}_s^{12} \\ \mathbf{J}_s^{21} & \mathbf{J}_s^{22} \end{pmatrix}, \tag{5.28}$$

where

$$\begin{aligned} \mathbf{J}_s^{11} &= \mathbf{J}_s(1,1), & (5.29) \\ \mathbf{J}_s^{12} &= \mathbf{J}_s^{21\,H} \\ &= [\mathbf{J}_s(1,2)\ \mathbf{J}_s(1,3)\ \ldots\ \mathbf{J}_s(1,K+1)], & (5.30) \\ \mathbf{J}_s^{22} &= \mathrm{diag}\left[\mathbf{J}_s(2,2)\ \mathbf{J}_s(3,3)\ \cdots\ \mathbf{J}_s(K+1,K+1)\right], & (5.31) \end{aligned}$$

where "diag[ ]" represents a diagonal matrix.

For the 2×2 matrix (5.28), it is convenient to represent the (1,1) element of its inverse matrix as:

$$\mathbf{J}_s^{-1}(1,1) = \left(\mathbf{J}_s^{11} - \mathbf{J}_s^{12}\left(\mathbf{J}_s^{22}\right)^{-1}\mathbf{J}_s^{21}\right)^{-1}. \tag{5.32}$$

$\mathbf{J}_s^{22}$ is a diagonal matrix, hence its inverse, $\left(\mathbf{J}_s^{22}\right)^{-1}$, is also diagonal with each element the reciprocal of the original $\mathbf{J}_s^{22}$ diagonal element. Inserting (5.29) - (5.31) into (5.32) gives

$$\begin{aligned} \mathbf{CRB}(\epsilon) &= \mathbf{J}_s^{-1}(1,1) & (5.33) \\ &= \left(\mathbf{J}_s(1,1) - \sum_{k=1}^{K}\frac{|\mathbf{J}_s(1,k+1)|^2}{\mathbf{J}_s(k+1,k+1)}\right)^{-1}. \end{aligned}$$

Equation (5.26) indicates that the first and second derivatives of the conditional log-likelihood function with respect to each unknown parameter must be determined first.

Start with the first derivative of ln p$\{\mathbf{R}|\vec{\Theta}_s\}$ with respect to ϵ to give

$$\frac{\partial \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial \epsilon} = -\sum_{k=1}^{K}(\vec{R}_k - \vec{S}_k)^H \Lambda_k^{-1}\frac{\partial(-\vec{S}_k)}{\partial \epsilon}. \tag{5.34}$$

Then take the second derivative of ln p$\{\mathbf{R}|\vec{\Theta}_s\}$ with respect to ϵ to give

$$\begin{aligned} &\frac{\partial^2 \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial \epsilon^2} \\ &= -\sum_{k=1}^{K}\left[\frac{\partial(-\vec{S}_k)^H}{\partial \epsilon}\Lambda_k^{-1}\frac{\partial(-\vec{S}_k)}{\partial \epsilon}\right. \\ &\left.+(\vec{R}_k - \vec{S}_k)^H\Lambda_k^{-1}\frac{\partial^2(-\vec{S}_k)}{\partial \epsilon^2}\right]. \end{aligned} \tag{5.35}$$

Given (5.16), (5.18), and (5.26), and then taking the expected value of (5.35), (5.29) can be written as

$$\begin{aligned} \mathbf{J}_s^{11} = \mathbf{J}_s(1,1) &= -E\left[\frac{\partial^2 \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial \epsilon^2}\right] \\ &= \sum_{k=1}^{K}\frac{\partial \vec{S}_k^H}{\partial \epsilon}\Lambda_k^{-1}\frac{\partial \vec{S}_k}{\partial \epsilon}. \end{aligned} \tag{5.36}$$

Next evaluate $\mathbf{J}_s^{12}$ and $\mathbf{J}_s^{21}$. Taking the derivative of (5.34) with respect to s_k, and then taking the expected value, the elements of $\mathbf{J}_s^{12}$ and $\mathbf{J}_s^{21}$ are given by

$$\begin{aligned}
\mathbf{J}_s(1,k+1) &= \mathbf{J}_s(k+1,1) \\
&= -E\left[\frac{\partial^2 \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial\vec{\Theta}_s(1)\partial\vec{\Theta}_s(k+1)}\right] \\
&= -E\left[\frac{\partial^2 \ln \mathrm{p}\{\mathbf{R}|\vec{\Theta}_s\}}{\partial\epsilon\partial s_k}\right] \\
&= \frac{\partial\vec{S}_k^H}{\partial s_k}\Lambda_k^{-1}\frac{\partial\vec{S}_k}{\partial\epsilon}. \qquad (5.37)
\end{aligned}$$

Similarly, the $\mathbf{J}_s^{22}$ elements represented by (5.27) become

$$\begin{aligned}
&\mathbf{J}_s(i+1,k+1) \\
&= \begin{cases} \frac{\partial\vec{S}_k^H}{\partial s_k}\Lambda_k^{-1}\frac{\partial\vec{S}_k}{\partial s_k}, & i=k,\ 1\le i,k\le K; \\ 0, & i\ne k,\ 1\le i,k\le K. \end{cases} \qquad (5.38)
\end{aligned}$$

If $s_0(t)$ is known, the time delayed and attenuated source signal at each sensor is known. Hence the unknown parameter vector $\vec{\Theta}_s$ only needs to include $\{s_k\}$ $(k = 1, \cdots, K)$, denoting the components of the original source signal $s_0(t)$. In contrast, the relationship between s_k and $\vec{S}_k$ is not as clear as that of $s_0(t)$ and $s_m(t)$ in the time domain. As everything now has to be manipulated in the domain defined by $\{\phi_k(t)\}$ $(k = 1, \cdots, K)$, it is necessary to find out how s_k and $\vec{S}_k$ are related to each other.

Perform the element-wise division, denoted "./", on both sides of (5.11) by $\vec{d} = [d_1\ d_2\ \cdots\ d_M]^T$ and replace t with

$$\begin{aligned}
\vec{t} &= [t_1\ t_2\ \cdots\ t_M]^T \\
&= [t-\tau_1\ t-\tau_2\ \cdots\ t-\tau_M]^T, \qquad (5.39)
\end{aligned}$$

to give

$$\vec{S}(t) = \sum_{k=1}^{K\to\infty} s_k\phi_k(\vec{t})^*./\,\vec{d}. \qquad (5.40)$$

Comparing (5.40) with (5.13), it is seen that the left sides of both equations are equal, which means the right sides of both must also be equal. Thus

$$\vec{S}_k\phi_k(t)^* = s_k\phi_k(\vec{t})^*./\,\vec{d}, \qquad (5.41)$$

which gives

$$\frac{\partial\vec{S}_k}{\partial s_k} = \phi_k(\vec{t})^*./\,[\phi_k(t)^*\vec{d}]. \qquad (5.42)$$

With both $\vec{d}$ and $\phi_k(\vec{t})$ being functions of ϵ, the chain rule for composite functions can be applied to (5.41) to derive $\frac{\partial \vec{S}_k}{\partial \epsilon}$, which gives

$$\frac{\partial \vec{S}_k}{\partial \epsilon} = \frac{s_k}{\phi_k(t)^*} \frac{\partial \vec{d}}{\partial \epsilon}. \times \left[\phi_k'(\vec{t})^* ./ (c\vec{d}) - \phi_k(\vec{t})^* ./ (\vec{d}.^2) \right], \tag{5.43}$$

where

$$\phi'(\vec{t})^* = d\phi(\vec{t})^* ./ d\vec{t}, \tag{5.44}$$

$$\frac{\partial \vec{d}}{\partial \epsilon} = [\frac{\partial d_1}{\partial \epsilon} \ \frac{\partial d_2}{\partial \epsilon} \ \cdots \frac{\partial d_M}{\partial \epsilon}]^T, \tag{5.45}$$

with

$$\frac{\partial d_m}{\partial \epsilon} = -\frac{\vec{e} \cdot \vec{p}_m}{d_m}, \tag{5.46}$$

".×" and ".^{2}" denoting the element-wise multiplication and the element-wise square.

Now substitute (5.42) and (5.43) back into (5.36) - (5.38) to give

$$\begin{aligned}
&\mathbf{J}_s(1,1) \\
= \ & \sum_{k=1}^{K} \sum_{m=1}^{M} \Lambda_k^{-1}(m,m) \left| \frac{s_k}{\phi_k(t)} \right|^2 \left(\frac{\partial d_m}{\partial \epsilon} \right)^2 \\
& \times \left[\frac{|\phi_k'(t_m)|^2}{c^2 d_m^2} - \frac{\phi_k'(t_m)^* \phi_k(t_m)}{c d_m^3} \right. \\
& \left. - \frac{\phi_k'(t_m) \phi_k(t_m)^*}{c d_m^3} + \frac{|\phi_k(t_m)|^2}{d_m^4} \right] \\
& + 2 \sum_{k=1}^{K} \sum_{m=1}^{M} \sum_{n=m+1}^{M} \Lambda_k^{-1}(m,n) \left| \frac{s_k}{\phi_k(t)} \right|^2 \frac{\partial d_m}{\partial \epsilon} \frac{\partial d_n}{\partial \epsilon} \\
& \times \left[\frac{\phi_k'(t_m)^* \phi_k'(t_n)}{c^2 d_m d_n} - \frac{\phi_k'(t_m)^* \phi_k(t_n)}{c d_m^2 d_n} \right. \\
& \left. - \frac{\phi_k'(t_m) \phi_k(t_n)^*}{c d_m d_n^2} + \frac{\phi_k(t_m) \phi_k(t_n)^*}{d_m^2 d_n^2} \right],
\end{aligned} \tag{5.47}$$

$$\begin{aligned}
&\mathbf{J}_s(1,k+1) \\
= \ & \sum_{m=1}^{M} \frac{s_k \phi_k(t_m) \Lambda_k^{-1}(m,m)}{|\phi_k(t)|^2 d_m} \frac{\partial d_m}{\partial \epsilon} \left[\frac{\phi_k'(t_m)^*}{c d_m} - \frac{\phi_k(t_m)^*}{d_m^2} \right] \\
& + 2 \sum_{m=1}^{M} \sum_{n=m+1}^{M} \frac{s_k \phi_k(t_n) \Lambda_k^{-1}(m,n)}{|\phi_k(t)|^2 d_n} \frac{\partial d_m}{\partial \epsilon} \\
& \qquad \times \left[\frac{\phi_k'(t_m)^*}{c d_m} - \frac{\phi_k(t_m)^*}{d_m^2} \right],
\end{aligned} \tag{5.48}$$

$$
\begin{aligned}
&\mathbf{J}_s(i+1,k+1) \\
&= \begin{cases}
\sum_{m=1}^{M} \Lambda_k^{-1}(m,m) \left| \frac{\phi_k(t_m)}{\phi_k(t) d_m} \right|^2 \\
+2 \sum_{m=1}^{M} \sum_{n=m+1}^{M} \Lambda_k^{-1}(m,n) \\
\quad \times \frac{\phi_k(t_m)\phi_k(t_n)^*}{\phi_k(t)\phi_k(t)^* d_m d_n}, & i = k; \\
\\
0, & i \neq k,
\end{cases}
\end{aligned} \tag{5.49}
$$

where $1 \leq i, k \leq K$.

If the class of functions for $\{\phi_k(t)\}$ $(k = 1, \cdots, K)$ satisfy the following features,

$$
\begin{aligned}
\phi_k(a+b) &= \phi_k(a)\phi_k(b), && (5.50) \\
\frac{\partial \phi_k(c_k a)}{\partial a} &= c_k \phi_k(a), && (5.51)
\end{aligned}
$$

where a, b, and c_k are arbitrary complex values, (5.42) and (5.43) become

$$
\begin{aligned}
\frac{\partial \vec{S}_k}{\partial s_k} &= \phi_k(\vec{\tau})./\vec{d}, && (5.52) \\
\frac{\partial \vec{S}_k}{\partial \epsilon} &= s_k \frac{\partial \vec{d}}{\partial \epsilon}. \times \left[c_k \phi_k'(\vec{\tau})./(c\vec{d}) - \phi_k(\vec{\tau})./(\vec{d}.^2) \right], && (5.53)
\end{aligned}
$$

where $\vec{\tau} = [\tau_1 \ \ \tau_2 \ \ \cdots \tau_M]^T$. Hence (5.47), (5.48), and (5.49) are reduced to

$$
\begin{aligned}
&\mathbf{J}_s(1,1) \\
&= \sum_{k=1}^{K} \sum_{m=1}^{M} |s_k \phi_k(\tau_m)|^2 \Lambda_k^{-1}(m,m) \left(\frac{\partial d_m}{\partial \epsilon} \right)^2 \\
&\quad \times \left[\frac{|c_k|^2}{c^2 d_m^2} - \frac{c_k^*}{c d_m^3} - \frac{c_k}{c d_m^3} + \frac{1}{d_m^4} \right] \\
&+ \ 2 \sum_{m=1}^{M} \sum_{n=m+1}^{M} |s_k|^2 \phi_k(\tau_m)\phi_k(\tau_n)^* \Lambda_k^{-1}(m,n) \frac{\partial d_m}{\partial \epsilon} \\
&\quad \times \frac{\partial d_n}{\partial \epsilon} \left[\frac{|c_k|^2}{c^2 d_m d_n} - \frac{c_k^*}{c d_m^2 d_n} - \frac{c_k}{c d_m d_n^2} + \frac{1}{d_m^2 d_n^2} \right],
\end{aligned} \tag{5.54}
$$

$$
\begin{aligned}
&\mathbf{J}_s(1,k+1) \\
&= \sum_{m=1}^{M} \frac{s_k |\phi_k(\tau_m)|^2 \Lambda_k^{-1}(m,m)}{d_m} \frac{\partial d_m}{\partial \epsilon} \left[\frac{c_k}{c d_m} - \frac{1}{d_m^2} \right] \\
&\quad +2 \sum_{m=1}^{M} \sum_{n=m+1}^{M} \frac{s_k \phi_k(\tau_m)\phi_k(\tau_n)^* \Lambda_k^{-1}(m,n)}{d_n} \\
&\quad \times \frac{\partial d_m}{\partial \epsilon} \left[\frac{c_k}{c d_m} - \frac{1}{d_m^2} \right],
\end{aligned} \tag{5.55}
$$

$$\mathbf{J}_s(i+1,k+1) = \begin{cases} \sum_{m=1}^{M} \Lambda_k^{-1}(m,m) \left| \frac{\phi_k(\tau_m)}{d_m} \right|^2 \\ +2\sum_{m=1}^{M}\sum_{n=m+1}^{M} \frac{\phi_k(\tau_m)\phi_k(\tau_n)^*}{d_m d_n} \\ \times \Lambda_k^{-1}(m,n), & i = k; \\ \\ 0, & i \neq k; \end{cases} \tag{5.56}$$

where $1 \leq i, k \leq K$.

In order to evaluate $\mathbf{CRB}(\epsilon)$, the basis function set $\{\phi_k(t)\}$ $(k = 1, \cdots, K)$ has to be specified. In most signal processing applications, it is preferred to analyze signals in the frequency domain using the well-known Fourier techniques. However, the two-dimensional Fourier transform of a non-stationary kernel results in a function of two frequency variables. This indicates that the basis function set $\{e^{j2\pi ft}\}$, where f is a continuous real valued variable, is not a simple direct solution to the integral equation given by (5.8). Instead, some other closed-form solution is needed. For positive definite kernels, as given in the matrix of covariance functions of random noise processes denoted $\mathbf{C}_{NS}(t_1, t_2)$, Mercer's theorem ensures that there exists a certain set of eigenfunctions that can expand the kernels bilinearly with uniform convergence. This means that the corresponding eigenvalues form the desired uncorrelated coefficients in the new coordinate system represented by these eigenfunctions, and analytic CRB expressions shall be achievable in this new system. However, derivation of possible analytic CRB expressions for non-stationary processes is left for future work to limit the length of this paper.

The focus of the following context is on kernels that satisfy the special feature of $\mathbf{C}_{NS}(t_2 - t_1) = \mathbf{C}_{NS}(\tau)$. The matrix of covariance functions, $\mathbf{C}_{NS}(t_2 - t_1)$, stands for a group of stationary processes, for which the basis function set $\{e^{j2\pi ft}\}$, $f \in R$ is known [23] to satisfy the integral equation (5.8) and the additional requirements defined by (5.50) and (5.51) at the same time. However, this classical FT performed over continuous signals gives continuous eigenvalues, which are not desirable for the construction of the log-likelihood function. Thus a modified set of basis functions that still satisfies the requirements is desired to give discrete eigenvalues. A natural choice to try is

$$\left\{\phi_k(t) = e^{-j2\pi f_k t}\right\}, \qquad k = 1, 2, \cdots, K., \tag{5.57}$$

where the difference between f_k and f_{k+1} is a constant Δf Hz. It is apparent that (5.57) asymptotically approaches $\{e^{j2\pi ft}\}$ when $\Delta f \to 0$, and it is a solution to (5.8).

It is obvious that

$$\begin{aligned} \phi_k(t_1 - t_2) &= e^{-j2\pi f_k t_2} e^{+j2\pi f_k t_1}, & (5.58) \\ c_k &= -j2\pi f_k, & (5.59) \\ c_k^* + c_k &= 0. & (5.60) \end{aligned}$$

Examining the source signal, it can be expanded and written explicitly as

$$s_0(t) = \sum_{\substack{k=1 \\ \Delta f \to 0}}^{K \to \infty} s_k e^{j2\pi f_k t}, \tag{5.61}$$

where

$$s_k = \int_{-\infty}^{\infty} s_0(t) e^{-j2\pi f_k t} dt. \tag{5.62}$$

Though (5.57) produces a solution to (5.8), the completeness remains a question. Let $X_s(f)$ denote the classical continuous Fourier transform of $s_0(t)$,

$$s_0(t) = \int_{-\infty}^{\infty} X_s(f) e^{j2\pi f t} df, \tag{5.63}$$

which can also be represented by a sum of infinitesimal increments

$$s_0(t) = \sum_{\substack{k=1 \\ \Delta f \to 0}}^{K \to \infty} X_s(f_k) e^{j2\pi f_k t} \Delta f. \tag{5.64}$$

Comparing the right sides of (5.61) and (5.64) gives

$$s_k = \Delta f X_s(f_k). \tag{5.65}$$

This step ensures that the energy in the original source signal is conserved and thus the condition of completeness is met.

Note that if there is only a finite number of K basis functions given a certain Δf, the accuracy of approximating a waveform by the group of $\left\{e^{-j2\pi f_k t}\right\}$ $(k = 1, \cdots, K)$ is affected by the smoothness of the waveform envelope in the frequency domain. The truncation error can be large.

Examining the noise, let $S_{NS}^{mn}(f)$ denote the result of applying the classical continuous FT to the covariance $\mathbf{C}_{NS}^{mn}(\tau)$ of the random noise processes at sensor m and sensor n,

$$S_{NS}^{mn}(f) = \int_{-\infty}^{\infty} \mathbf{C}_{NS}^{mn}(\tau) e^{-j2\pi f \tau} d\tau. \tag{5.66}$$

The inverse transform is given by

$$\mathbf{C}_{NS}^{mn}(\tau) = \int_{-\infty}^{\infty} S_{NS}^{mn}(f) e^{j2\pi f \tau} df, \tag{5.67}$$

and its representation by a sum of infinitesimal increments is

$$\mathbf{C}_{NS}^{mn}(\tau) = \sum_{\substack{k=1 \\ \Delta f \to\ 0}}^{K\to\infty} S_{NS}^{mn}(f_k) e^{j2\pi f_k \tau} \Delta f. \tag{5.68}$$

Now refer back to (5.9). Given (5.57), (5.68) can be also written as

$$\mathbf{C}_{NS}^{mn}(t_1, t_2) = \sum_{\substack{k=1 \\ \Delta f \to\ 0}}^{K\to\infty} \Lambda_k(m,n) e^{-j2\pi f_k t_1} e^{j2\pi f_k t_2}. \tag{5.69}$$

Similarly, in order to ensure the completeness in the expansion using (5.57), (5.69) needs to be compared with (5.68). Given $\tau = t_2 - t_1$ and $\mathbf{C}_{NS}^{mn}(t_1, t_2) = \mathbf{C}_{NS}^{mn}(\tau)$, it is seen that

$$\Lambda_k(m,n) = \Delta f S_{NS}^{mn}(f_k). \tag{5.70}$$

Obviously when Δf approaches zero, the results given in (5.65) and (5.70) are identical to those obtained after performing the integrated Fourier transforms of the corresponding input, as shown in [23] .

By now, the CRB for spatially correlated stationary CGN can be calculated trivially for certain signal and noise given in the forms of $X_s(f_k)$ and $S_{NS}^{mn}(f_k)$. Next, the derivation leading to more compact CRB expressions for spatially uncorrelated stationary Gaussian noise cases is presented in terms of continuous signal energy spectrum and noise power spectrum.

In many source localization applications, the sensor placement is sparse. This implies that the correlation between the sensors can be neglected and accordingly, Λ_k becomes a diagonal matrix given by

$$\Lambda_k = \mathrm{diag}[\lambda_{k1} \;\; \lambda_{k2} \;\; \cdots \lambda_{kM}], \tag{5.71}$$

and accordingly

$$S_{NS}^{mn} = \begin{cases} 0, & m \neq n; \\ S_{NS}^{m}, & m = n. \end{cases} \tag{5.72}$$

With (5.60), (5.65), (5.70), and (5.72), (5.54) becomes

$$\begin{aligned} & \mathbf{J}_s(1,1) \\ = \; & \sum_{m=1}^{M} \left(\frac{\partial d_m}{\partial \epsilon} \right)^2 \sum_{k=1}^{K\to\infty} \frac{|X_s(f_k)\Delta f|^2}{S_{NS}^{m}(f_k)\Delta f} \times \\ & \left| \frac{(2\pi f_k)^2}{c^2 d_m^2} + \frac{1}{d_m^4} \right|. \end{aligned} \tag{5.73}$$

To determine (5.33), $|\mathbf{J}_s(1,k)|^2$ has to be found. Square $|\mathbf{J}_s(1,k)|$ given by (5.55) and then substitute

with (5.57) to give

$$\begin{aligned} & |\mathbf{J}_s(1,k+1)|^2 \\ = & \sum_{m=1}^{M}\sum_{n=1}^{M} \frac{|s_k|^2}{\lambda_{km}\lambda_{kn}} \frac{\partial d_m}{\partial \epsilon}\frac{\partial d_n}{\partial \epsilon} \times \\ & \left[\frac{c_k c_k^*}{c^2 d_m^2 d_n^2} - \frac{c_k}{c d_n^2 d_m^3} - \frac{c_k^*}{c d_m^2 d_n^3} + \frac{1}{d_m^3 d_n^3}\right]. \end{aligned} \tag{5.74}$$

Substituting (5.60), (5.65), (5.70), and (5.72) into (5.74) gives

$$\begin{aligned} & |\mathbf{J}_s(1,k+1)|^2 \\ = & \sum_{m=1}^{M}\sum_{n=1}^{M} \frac{|X_s(f_k)\Delta f|^2}{S_{NS}^m(f_k)\Delta f S_{NS}^n(f_k)\Delta f} \times \\ & \frac{\partial d_m}{\partial \epsilon}\frac{\partial d_n}{\partial \epsilon}\left[\frac{(2\pi f_k)^2}{c^2 d_m^2 d_n^2} + \frac{1}{d_m^3 d_n^3}\right], \end{aligned} \tag{5.75}$$

since

$$\sum_{m=1}^{M}\sum_{n=1}^{M} -\frac{j2\pi f_k}{c d_n^2 d_m^3} - \frac{-j2\pi f_k}{c d_m^2 d_n^3} = 0. \tag{5.76}$$

Substituting (5.57) , (5.70), and (5.72) into (5.56) gives

$$\begin{aligned} & \mathbf{J}_s(i+1,k+1) \\ = & \begin{cases} \sum_{m=1}^{M} \frac{1}{d_m^2 S_{NS}^m(f_k)\Delta f}, & i=k, 1\le i,k\le K; \\ 0, & i\ne k,\ 1\le i,k\le K. \end{cases} \end{aligned} \tag{5.77}$$

Hence,

$$\begin{aligned} & \sum_{k=1}^{K} \frac{|\mathbf{J}_s(1,k+1)|^2}{\mathbf{J}_s(k+1,k+1)} \\ = & \left\{\sum_{k=1}^{K}\left[\sum_{m=1}^{M}\sum_{n=1}^{M} \frac{|X_s(f_k)\Delta f|^2}{S_{NS}^m(f_k)\Delta f S_{NS}^n(f_k)\Delta f} \times \right.\right. \\ & \left.\left. \frac{\partial d_m}{\partial \epsilon}\frac{\partial d_n}{\partial \epsilon}\left(\frac{(2\pi f_k)^2}{c^2 d_m^2 d_n^2} + \frac{1}{d_m^3 d_n^3}\right)\right]\right\} \\ & \Big/ \sum_{m=1}^{M} \frac{1}{d_m^2 S_{NS}^m(f_k)\Delta f}\ . \end{aligned} \tag{5.78}$$

Substituting (5.46) into (5.73) and then letting $\Delta f \to 0$ and $K \to \infty$ gives a continuous function of the signal energy spectral density and noise power spectral density in the frequency domain

$$\begin{aligned} & \mathbf{J}_s(1,1) \\ = & \frac{1}{2\pi}\int_{-\Omega_B}^{\Omega_B} \sum_{m=1}^{M} q_m^2 \frac{S_{SIG}(\omega)}{S_{NS}^m(\omega)}\left(\frac{\omega^2}{c^2} + \frac{1}{d_m^2}\right) d\omega, \end{aligned} \tag{5.79}$$

where $\omega = 2\pi f$, $S_{SIG}(\omega) = |X_s(\omega)|^2$ is the energy spectrum of $s(t)$ with a finite bandwidth Ω_B in radians, $S_{NS}(\omega)$ is the power spectrum of the noise and

$$q_m = -\frac{\vec{e}\cdot\vec{p}_m}{d_m^2}. \tag{5.80}$$

Similarly, (5.78) can be expressed as

$$\begin{aligned}
&\sum_{k=2}^{K} \frac{|\mathbf{J}_s(1,k+1)|^2}{\mathbf{J}_s(k+1,k+1)} \\
= \quad & \frac{1}{2\pi}\int_{-\Omega_B}^{\Omega_B} \left\{ \sum_{m=1}^{M}\sum_{n=1}^{M} \frac{S_{SIG}(\omega)}{S_{NS}^m(\omega)S_{NS}^n(\omega)} \frac{q_m q_n}{d_m d_n} \times \right. \\
& \left. \left[\frac{\omega^2}{c^2} + \frac{1}{d_m d_n}\right] \right\} \Bigg/ \sum_{m=1}^{M} \frac{1}{d_m^2 S_{NS}^m(\omega)} d\omega \ .
\end{aligned} \tag{5.81}$$

Substituting (5.79) and (5.81) into (5.33) gives the CRB expression, $\mathbf{CRB}(\epsilon)_{CGN-dif}$ in (5.82). Here the noise is stationary coloured Gaussian with a different power spectrum at each sensor. It is also spatially uncorrelated from sensor to sensor.

$$\begin{aligned}
&\mathbf{CRB}(\epsilon)_{CGN-dif} = \\
&2\pi \left\{ \int_{-\Omega_B}^{\Omega_B} \sum_{m=1}^{M} q_m^2 \frac{S_{SIG}(\omega)}{S_{NS}^m(\omega)} \left(\frac{\omega^2}{c^2} + \frac{1}{d_m^2}\right) d\omega \right. \\
&- \int_{-\Omega_B}^{\Omega_B} \left[\sum_{m=1}^{M}\sum_{n=1}^{M} \frac{S_{SIG}(\omega)}{S_{NS}^m(\omega)S_{NS}^n(\omega)} \frac{q_m q_n}{d_m d_n} \times \right. \\
&\left. \left. \left(\frac{\omega^2}{c^2} + \frac{1}{d_m d_n}\right) \right] \Bigg/ \sum_{m=1}^{M} \frac{1}{d_m^2 S_{NS}^m(\omega)} d\omega \right\}^{-1} .
\end{aligned} \tag{5.82}$$

If the stationary CGN power spectrum at each sensor is equal and all denoted $S_{NS}(\omega)$, the CRB expression in (5.82) is reduced to

$$\begin{aligned}
&\mathbf{CRB}(\epsilon)_{CGN-sme} = \\
&2\pi \left\{ \int_{-\Omega_B}^{\Omega_B} \frac{\omega^2 S_{SIG}(\omega)}{c^2 S_{NS}(\omega)} d\omega \left[\sum_{m=1}^{M} q_m^2 - \frac{\left(\sum_{m=1}^{M} \frac{q_m}{d_m}\right)^2}{\sum_{m=1}^{M} \frac{1}{d_m^2}} \right] \right. \\
&\left. + \int_{-\Omega_B}^{\Omega_B} \frac{S_{SIG}(\omega)}{S_{NS}(\omega)} d\omega \left[\sum_{m=1}^{M} \frac{q_m^2}{d_m^2} - \frac{\left(\sum_{m=1}^{M} \frac{q_m}{d_m^2}\right)^2}{\sum_{m=1}^{M} \frac{1}{d_m^2}} \right] \right\}^{-1} .
\end{aligned} \tag{5.83}$$

Further simplification of the noise to identical WGN random process, where $S_{NS}(\omega)$ is flat and equals

$\frac{\pi\sigma^2}{\Omega_B}$ with σ^2 being the common variance of $\mathbf{n}_m(t)$, $m = 1, 2, \cdots, M$, in the time domain, gives

$$\begin{aligned}
&\mathbf{CRB}(\epsilon)_{WGN} \\
= &\left\{ \left[\sum_{m=1}^{M} q_m^2 - \frac{1}{\sum_{m=1}^{M} \frac{1}{d_m^2}} \left(\sum_{m=1}^{M} \frac{q_m}{d_m} \right)^2 \right] \times \right. \\
&\frac{\Omega_B \int_{-\Omega_B}^{\Omega_B} \omega^2 S_{SIG}(\omega) d\omega}{2\pi^2 \sigma^2 c^2} \\
&+ \left[\sum_{m=1}^{M} \frac{q_m^2}{d_m^2} - \frac{1}{\sum_{m=1}^{M} \frac{1}{d_m^2}} \left(\sum_{m=1}^{M} \frac{q_m}{d_m^2} \right)^2 \right] \times \\
&\left. \frac{\Omega_B \int_{-\Omega_B}^{\Omega_B} S_{SIG}(\omega) d\omega}{2\pi^2 \sigma^2} \right\}^{-1}. \qquad (5.84)
\end{aligned}$$

If a certain set of orthonormal basis functions can be found for a kernel representing a non-stationary Gaussian noise covariance, the continuous eigenvalues corresponding to the eigenfunctions may be defined as, e.g., $Q_s(\rho)$ and $Q_n(\rho)$ for the signal and noise, respectively. If this basis also satisfies (5.50) and (5.51), a CRB expression in the form of (5.82) may be reached for cases of non-stationary Gaussian noise using the signal and noise representation of $Q_s(\rho)$ and $Q_n(\rho)$ in the new domain.

5.4 Application of the CRB expressions

In this section, an example of using the derived CRB expressions for algorithm evaluation purpose is demonstrated. An energy-based grid search (EGS) localization algorithm [15] is chosen and Monte Carlo simulation is performed. The EGS algorithm estimates the source location with the hypothesized location that produces the greatest energy in the sum of sensor output signals. These sensor signals have been time-shifted by amounts consistent with the propagation delays of hypothesized locations of the event. The source location estimation error of the EGS algorithm calculated from the Monte Carlo simulation result can then be evaluated against the corresponding CRB expressions in the form of root mean squared error (RMSE).

The source signal used was recorded by a reference sensor placed in the Lanigan Mine in Saskatchewan, Canada. The DC bias has been removed, and the negligible noise is taken as part of the original source signal. The sensors are arbitrarily placed within the mine, and their coordinates are given. The CGN is generated by filtering white Gaussian noise. Since the source signal energy is concentrated in the frequency band of 50–200 Hz, the simulated sensor signals are assumed to have been filtered by a corresponding low pass filter. The signal propagation speed adopted for the simulation is 2581 m/s, which is the sound speed through rocks.

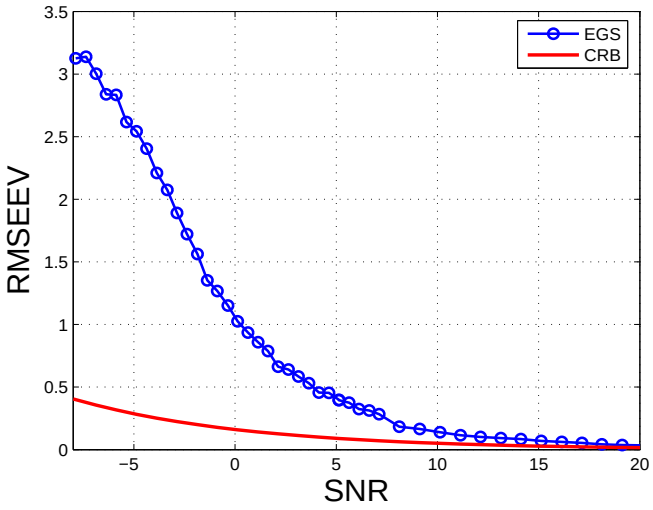

Figure 5.1: Monte Carlo simulation of the EGS algorithm vs. CRB for the case of spatially independent stationary CGN with an unequal covariance at each sensor.

Figure 5.1 shows that with the value of signal-to-noise ratio (SNR) increasing, the performance of the EGS localization algorithm for spatially uncorrelated stationary CGN cases approaches the theoretical CRB values calculated from (5.82).

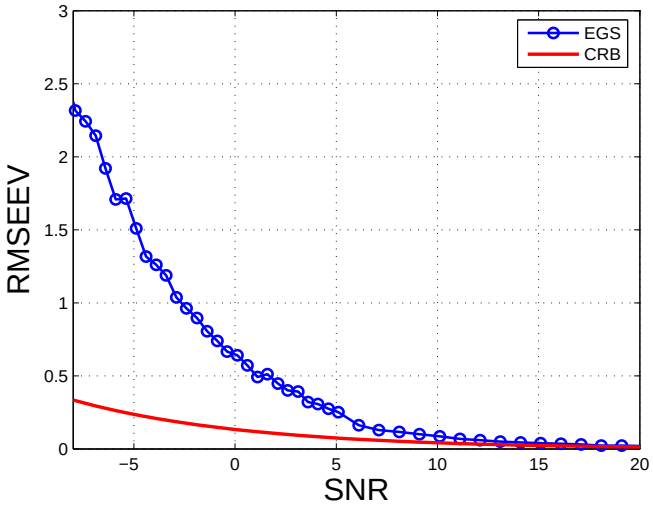

Figure 5.2: Monte Carlo simulation of the EGS algorithm vs. CRB for the case of WGN with an equal covariance at each sensor.

Figure 5.2 shows the correspondence between the EGS performance and the CRB calculated from (5.84)

for WGN cases. The same trends for the EGS and CRB curves are seen, as expected. The Monte Carlo simulation performed to generate Figure 5.1 and Figure 5.2 used 1000 iterations to produce the RMSE at each SNR value.

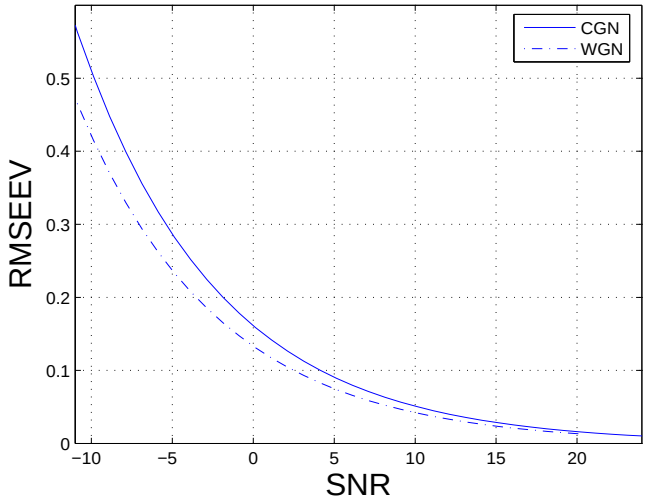

Figure 5.3: CRB plots for the two noise cases aforementioned, the spatially independent stationary CGN and WGN.

In Figure 5.3, the CRB curves for the two noise cases considered in the simulation are plotted together for easier comparison. It indicates that for the type of source signal and sensor configuration adopted for the simulation and calculation done in this paper, better estimation accuracy can be obtained in WGN environment than in CGN environment.

In either CGN or WGN scenarios, the EGS performance asymptotically converges to the CRB, which proves that the EGS is truly an unbiased estimator, as expected. Then by comparing the EGS performance curve with that of the CRB under low SNR, it is shown clearly that there is room for the EGS algorithm to be optimized. In summary, the CRB expressions derived in this paper are very useful in theoretically evaluating a class of localization algorithms, such as the EGS algorithm.

5.5 Conclusion

In this paper, Cramér-Rao Bound derivation for source localization in a given direction for various forms of Gaussian noise is presented with closed-form CRB expressions arrived for several specialized cases.

The derivation is based on continuous-time sensor signals observed, instead of their discrete-time samples. The mathematical technique [23] used is applicable to the more general estimation/detection problems

that may involve coloured or even non-stationary Gaussian noise. Few researchers have considered this approach since a common idea is to consider only WGN, which significantly simplifies the mathematical issues. However, there are cases for which this model is not suitable, as discussed earlier. Though not as efficient as for WGN, discrete-time samples can still be used for stationary CGN. An additional "whitening" process will be needed to remove the correlation between the samples. It can be realized through the eigenvalue decomposition, a special case of the more general singular value decomposition [31]. Complex matrix calculation is required to find the corresponding orthogonal matrix that can diagonalize the covariance matrix of a CGN random process. Therefore, this paper deals with continuous signals directly. Doing so also makes it consistent with [15], which motivates this work.

To give an example of the application of the mathematical result presented in this paper, Monte Carlo simulation of the energy-based grid search algorithm has been performed for several specialized noise cases under study and then evaluated using the corresponding CRB expressions derived in this work.

Acknowledgments

The authors would like to thank TRlabs, Saskatoon and the Potash Corporation of Saskatchewan Inc., for supporting this research.

References

[1] E. by Y.T. Chan, "Underwater acoustic data processing," NATO ASI Seriers, Kluwer Academic Publishers, 1989.

[2] M. Brandstein, J. Adcock, and H. Silverman, "A closed-form location estimator for use with room environment microphone arrays," *IEEE Trans. Speech, and Audio Processing*, vol. 5, no. 1, pp. 45–50, 1997.

[3] Y. Huang, J. Benesty, and G. Elko, "Passive acoustic source localization for video camera steering," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 2, pp. 902–912, 2000.

[4] A. F. Prugger and D. J. Gendzill, "Microearthquake location: a nonlinear approach that makes use of a simplex stepping procedure," *Bulletin of th Seismological Society of America*, vol. 78, pp. 799–815, April 1988.

[5] D. Schaff, G. Bokelmann, W. Ellsworth, E. Zanzerkia, F. Waldhauser, and G. Beroza, "Optimizing correlation techniques for improved earthquake location," *Bulletin Seismological Soc. Am.*, vol. 94, pp. 705–721, April 2004.

[6] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 24, pp. 320–327, August 1976.

[7] G. C. Carter, *Coherence and Time Delay Estimation.* IEEE Press, 1993.

[8] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 35, pp. 1223–1225, August 1987.

[9] J. O. Smith and J. S. Abel, "The spherical interpolation method of source localization," *IEEE J. of Oceani. Eng.*, vol. 12, pp. 246–252, January 1987.

[10] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. on Signal Processing*, vol. 42, pp. 1905–1915, August 1994.

[11] Y. Huang, J. Benesty, G. Elko, and R. Mersereau, "Real-time passive source localization: a practical linear-correction least-squares approach," *IEEE Trans. on Speech and Audio Processing*, vol. 9, pp. 943–956, November 2001.

[12] H. Ye and R. D. Degroat, "Maximum likelihood DOA estimation and asymptotic Cramér-Rao bounds for additive unknown colored noise," *IEEE Trans. on Signal Processing*, vol. 43, pp. 938–949, April 1996.

[13] J. P. Delmas and H. Abeida, "Stochastic Cramér-Rao bound for noncircular signals with application to DOA estimation," *IEEE Trans. on Signal Processing*, vol. 52, pp. 3192–3199, November 2004.

[14] H. Abeida and J.-P. Delmas, "Gaussian Cramér-Rao bound for direction estimation of noncircular signals in unknown noise fields," *IEEE Trans. on Signal Processing*, vol. 53, pp. 4610–4618, December 2005.

[15] B. L. Daku and E. J. Salt, "Error analysis of a localization algorithm for finite-duration events," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1024–1034, March 2007.

[16] F. Gini, M. Luise, and R. Reggiannini, "Cramér-Rao bounds in the parametric estimation of fading radiotransmission channels," *IEEE Trans. on communications*, vol. 46, pp. 1390–1398, October 1998.

[17] F. Gini and G. B. Giannakis, "Hybrid FM-polynomial phase signal modeling: Parameter estimation and Cramér-Rao bounds," *IEEE Trans. on Signal Processing*, vol. 47, pp. 363–377, February 1999.

[18] M. Ghogho, A. K. Nandi, and A. Swami, "Cramér-Rao bounds and maximum likelihood estimation for random amplitude phase-modulated signals," *IEEE Trans. on Signal Processing*, vol. 47, pp. 2905–2916, November 1999.

[19] J. M. Francos and B. Friedlander, "Bounds for estimation of complex exponentials in unknown colored noise," *IEEE Trans. on Signal Processing*, vol. 43, pp. 2176–2185, September 1995.

[20] D. N. Swingler, "Approximate bounds on frequency estimates for short cisoids in colored noise," *IEEE Trans. on Signal Processing*, vol. 46, pp. 1456–1458, May 1998.

[21] K.W.Cheung and H.C.So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE Trans. on Signal Processing*, vol. 53, pp. 460–470, February 2005.

[22] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. on Signal Processing*, vol. 50, pp. 1843–1854, August 2002.

[23] H. L. V. Trees, *Detection, Estimation, and Modulation Theory.* John Wiley and Sons Inc., New York, 2001.

[24] M. Savic, M. Chugani, K. Peabody, A. Husain, M. Tang, and Z. Macek, "Digital signal processing aids cholesterol plaque detection," in *Proc. ICASSP'95*, (Detroit, MI), pp. 1173–1176, May 1995.

[25] R. J. Sclabassi, M. Sun, D. N. Krieger, P. Jasiukaitis, and M. S. Scher, "Time-frequency domain problems in the neurosciences," in *Time-Frequency Signal Analysis: Methods and Applications*, pp. 498–519, B. Boashash, Ed. and Longman-Cheshire Wiley Halsted Press, England, 1992.

[26] R. J. Sclabassi, M. Sun, D. N. Krieger, P. Jasiukaitis, and M. S. Scher, "Time-frequency analysis of the EEG signal," in *Proc. ISSP'90, Signal Processing, Theories, Implementations and Applications*, (Gold Coast, Australia), pp. 935–942, 1990.

[27] T. P. Wang, M. Sun, C. C. Li, and A. H. Vagnucci, "Classification of abnormal cortisol patterns by features from wigner spectra," in *Proc.of 10th International Conference on Pattern Recognition*, (Atlantic City, NJ), pp. 228–230, June 1990.

[28] D. E. Newland and G. D. Butler, "Application of time-frequency analysis to transient data from centrifuge earthquake testing," *Shock and Vibration*, vol. 7, pp. 195–202, 2000.

[29] W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. on Information Theory*, vol. IT-19, pp. 508–614, September 1973.

[30] S. T. Q. S. P. Huang and K. K. Phoon, "Convergence study of the truncated karhunen-loeve expansion for simulation of stochastic processes," *International Journal For Numerical Methods in Engineering*, vol. 52, pp. 1029–1043, 2001.

[31] J. M. Ortega, *Matrix Theory: A second course*. Plenum Press, New York, 1987.

6 Amplitude Weighting for Near-field Passive Source Localization

The content was originally published as:

S. Li and B. L. Daku, Optimal Amplitude Weighting for Near-field Passive Source Localization, IEEE Trans. on Signal Processing, vol. 59, pp. 6175–6185, August 2011.

The previous chapter presented the CRB for the class of estimators that uses the sum of cross-correlation between the sensor signals. The derivation is mathematically applicable to non-stationary Gaussian noise problems. However, closed-form expressions have been developed only for stationary Gaussian noise, either white or coloured.

The EGS algorithm, evaluated using the CRB obtained in the previous chapter, is shown that it can use some improvement. Therefore, this chapter proposes an amplitude weighting strategy that takes into account the difference between the sensor signal strengths when the EGS algorithm is applied to locate the source. A new way of implementing the amplitude weighted EGS algorithm in order to avoid additional calculations is also proposed.

For performance evaluation purposes, stationary Gaussian noise is adopted in this chapter. Note though that the amplitude weighted EGS algorithm is also applicable to non-stationary Gaussian noise cases, because no frequency components of the source signal or the noise plays a role in the algorithm.

Optimal Amplitude Weighting for Near-field Passive Source Localization

Sha Li* and Brian L.F. Daku, *Member, IEEE*

Abstract

In this paper, an optimal amplitude weight expression is derived and presented in closed form. The minimal error variance for location estimation of a near-field source signal that has a low time-bandwidth product is also presented. A noniterative method to calculate and apply weights is proposed to yield more robust estimation results at a lower calculational cost when compared with the traditional iterative method. Besides a theoretical evaluation, the proposed algorithm is also verified through Monte Carlo simulation. The weight expression derived also optimizes the system signal-to-noise ratio (SNR), and therefore it can be applied to improve the performance of any estimator/detector that utilizes the energy in the sum of sensor output signals.

Manuscript received February 09, 2011; revised June 18, 2011; accepted August 09, 2011. Date of publication August 30, 2011; date of current version November 16, 2011. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrick Flandrin. This work was supported by TRlabs, Saskatoon, and the Potash Corporation of Saskatchewan Inc.

The authors are with the Department of Electrical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: sha.li@usask.ca; brian.daku@usask.ca).

Digital Object Identifier10.1109/TSP.2011.2166393

Index Terms

Acoustic source localization, grid search, location estimation, microseismic.

6.1 Introduction

Passively locating an acoustic source is of great interest in a variety of applications, such as the tracking of underwater vehicles, video conferencing, mobile robot navigation and seismic event localization. Many traditional source location estimation techniques require a time difference of arrival (TDOA) estimation procedure [1] [2] since the source location is estimated with hyperbolic intersection or least square techniques [3–6], using the TDOA estimates. Such algorithms are referred to here as two-step algorithms and they assume that the acoustic signals emanated from various sources are either infinite or at least long enough to guarantee that the algorithms can perform as expected [7]. However, for a source that produces a short duration signal with a small product of the time duration and bandwidth, these two-step algorithms have difficulty performing above threshold, because they produce TDOA estimates with a larger error variance. Thus a more robust algorithm is needed for applications that involve localizing such a source.

One such algorithm is a single-step estimator utilizing an energy-based grid search (EGS) [8]. The comparison between the EGS algorithm, another one-step localization algorithm, and a classical two-step algorithm [9] indicates that the EGS yields the most robust performance and preferable for the application of interest. When the performance of the EGS algorithm is evaluated using the Cramér-Rao bound developed in [10], better estimation accuracy seems possible if some optimizing technique can be applied. Therefore, an amplitude weighted EGS (AWEGS) algorithm is proposed here. The primary contribution of this paper includes a closed-form amplitude weight expression derived using the error variance expression of the algorithm. And then the weights are applied to the EGS algorithm in a simple non-iterative way. The primitive idea of non-iterative weighting was first mentioned in the paper [11]. Further research was done afterwards, and this paper presents the latest results that are very important.

The specific application that motivated this study is microseismic event localization for safer underground mining. Microseismic events are usually human induced local earthquakes of low intensity, which are often resulted from rock rupture and shifting when the surrounding is under high stress caused by mining operations [8]. They may occur more than ten times per day within the locale of an underground mine, hence serve as a good indication of a potentially unstable area. Microseismic events usually last only a fraction of a second and are classified as near field occurrences, since monitoring sensors are normally deployed within the mine area. Monte Carlo simulation is used to evaluate the performance of the AWEGS algorithm using a real microseismic event signal recorded in the Lanigan potash mine in Saskatchewan, Canada. The performance

of the algorithm is also compared with the CRB from [10].

The paper is organized as follows: the system model used is described first, followed by a description of the AWEGS algorithm proposed and of the estimation error variance expression developed for it. Next, the signal-to-noise ratio (SNR) and error variance approaches to derive optimizing amplitude weights are investigated, and then the simple non-iterative way of applying the weights is introduced. Finally the original EGS and the proposed AWEGS algorithms are evaluated and compared, and brief conclusions are presented at the end.

6.2 System Model

In this paper, the system model is made as general as possible to be consistent with that of [10]. The model assumes the following: the signal is originated from a single-point source; the propagation medium is homogeneous and elastic (lossless); the propagation can be modeled as spherical spreading; any signals traveling along other than the direct path are sufficiently weak when they arrive at the sensors so that they can be ignored; the response of all the sensors are identical; and the output of each sensor is continuous and contains additive stationary Gaussian noise that is spatially independent with a zero-mean and common variance. All the sensor output signals are to be converted to lowpass signals by taking the magnitude of the analytic signal [12].

The spatial coordinate system used here is the three-dimensional x, y, z Cartesian system. The source position is given by $\vec{p}_s = (x_s, y_s, z_s)$. There are M sensors placed arbitrarily and sparsely in the system. The position of sensor m is denoted $\vec{p}_m = (x_m, y_m, z_m)$. The distance from the source to sensor m is denoted d_m and given by

$$\begin{aligned} d_m &= \| \vec{p}_m - \vec{p}_s \| \\ &= \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2 + (z_m - z_s)^2}. \end{aligned} \tag{6.1}$$

The output of sensor m, denoted $r_m(t)$, is a continuous sample function of a stochastic process $\mathbf{r}_m(t)$ that is given by

$$\mathbf{r}_m(t) = s_m(t) + \mathbf{n}_m(t), \tag{6.2}$$

where $\mathbf{n}_m(t)$ is an additive noise random process, and $s_m(t)$ is the signal received at sensor m that is independent of the noise. $s_m(t)$ is simply a delayed and attenuated version of the original unknown deterministic source signal $s_0(t)$ and is given by

$$s_m(t) = \frac{s_0(t - \tau_m)}{d_m}, \tag{6.3}$$

where $\tau_m = \frac{d_m}{c}$ is the time for the original signal $s_0(t)$ to travel from the source to sensor m, with c being the propagation speed of the mechanical wave. To accommodate this propagation model, it is assumed that

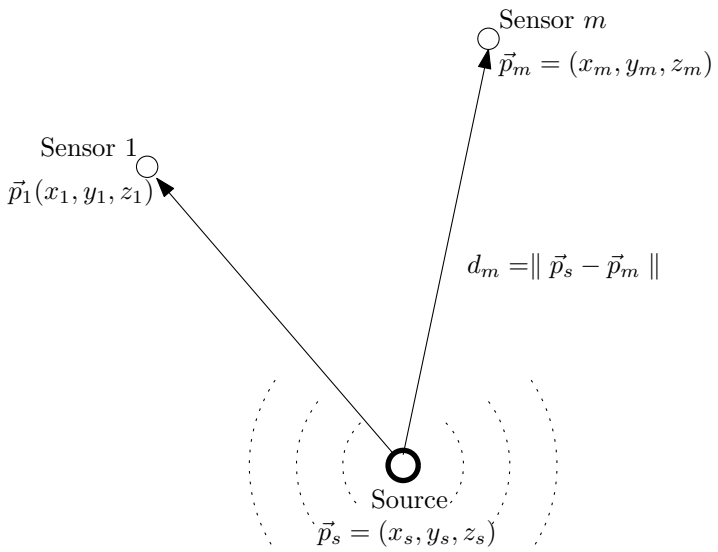

Figure 6.1: Coordinate system for the sensors

all the observed waveforms are collected by sensors that are at least one unit distance away from the source. The geometry of the sensor system is shown in Figure 6.1.

$s_0(t)$, denoting the real source signal, is assumed to be continuous with finite amplitude and a short-time duration T that can be estimated with negligible error at the closest sensor. The start time of the source event is taken as the time origin.

6.3 Proposed Amplitude Weighting

With each sensor signals amplitude being weighted, the original EGS algorithm becomes the amplitude weighted EGS algorithm. The idea proposed is to utilize the amplitude difference among the received signals, which is a reflection of the sensor-source distance. In this section, the weighted total energy function is introduced, followed by the presentation of the closed-form error variance expression derived for it. The SNR and error variance approaches to obtain weights are then investigated before the noniterative way to implement the AWEGS algorithm is presented.

6.3.1 Weighted Total Energy Function

The AWEGS method is based on the total energy function computed with the variable of the hypothesized source position, just as the EGS algorithm [8]. The difference is that each sensor signal is weighted, and hence a "weighted" total energy function results. The peak of this function indicates the estimate of the true source position. When optimal weights are applied, the peak becomes the sharpest, and the estimation

error variance will be minimal. The weighted total energy function is given by

$$W(\tilde{\vec{p}}_s, \vec{w}) = \int_0^T \left[\sum_{m=1}^{M} w_m r_m(t + \tilde{\tau}_m) \right]^2 dt, \tag{6.4}$$

where $\tilde{\vec{p}}_s$ is the hypothesized source position, w_m is the weight assigned to the m^{th} sensor, $\vec{w} = [w_1 \ w_2 \ \cdots \ w_M]$ the weight vector, and $\tilde{\tau}_m$, given by $\tilde{\tau}_m = \frac{\|\tilde{\vec{p}}_s - p_m\|}{c}$, is the time required for a mechanical wave to propagate from the hypothesized source position to sensor m. The limits of integration in practice are assumed to be estimated from the strongest sensor signal (i.e., the sensor closest to the event) and will have negligible error. It is seen that the original EGS algorithm [8]. is a special case of the AWEGS, where $w_m = 1$ for $m = 1, 2, \cdots, M$.

Since the noise at each sensor is assumed independent and identical Gaussian, and the grid search utilizes the phase difference between the delayed source signals across all the sensors, weights are applied to address only the amplitude difference. The goal is to weigh a stronger signal more than a weaker signal without changing the phase or frequency components of the signal during the localization process. Therefore, weight values should be positive and real, as weighting the amplitude with a negative or complex value does not help achieve the goal in this case. Thus weighting is applicable to any type of source signals.

6.3.2 Estimation Error Variance

An error analysis for the AWEGS estimator in a given direction in the presence of Gaussian noise as assumed was performed. Though the derivation is lengthy, a closed-form error expression for the AWEGS algorithm, denoted $\sigma_\epsilon^2(\vec{w})$, is reached after signal amplitude weighting is accommodated based on the work of [8].

Defining "E[ ]" to represent the expected value, and $\hat{\epsilon}$ the source position estimation error in a given direction denoted $\vec{e}$, there is

$$\sigma_\epsilon^2(\vec{w}) = E[\hat{\epsilon}^2] \approx (A + B)/D^2, \tag{6.5}$$

where

$$
\begin{aligned}
A &= I_A\Bigg[\sum_{k=1}^{M} w_k^2\Big(\sum_{m=1}^{M}\frac{w_m}{d_m}q_m\Big)^2 \\
&\quad -2\sum_{m=1}^{M}\frac{w_m}{d_m}q_m\sum_{k=1}^{M}w_k^2 q_k\sum_{l=1}^{M}\frac{w_l}{d_l} \\
&\quad +\Big(\sum_{m=1}^{M}\frac{w_m}{d_m}\Big)^2\sum_{k=1}^{M}(w_k q_k)^2\Bigg], \qquad (6.6)\\
B &= I_B\Bigg[\sum_{m=1}^{M} w_m^2\sum_{k=1}^{M}(w_k q_k)^2-\Big(\sum_{m=1}^{M}w_m^2 q_m\Big)^2\Bigg], \qquad (6.7)\\
D &= I_D\sum_{m=1}^{M}\sum_{k=1}^{M}\frac{w_m w_k}{d_m d_k}(q_m-q_k)^2, \qquad (6.8)
\end{aligned}
$$

and

$$
\begin{aligned}
I_A &= \frac{4}{2\pi c^2}\int_{-\infty}^{\infty}\omega^2 S_{ss}(\omega)S_{nn}(\omega)d\omega, \qquad (6.9)\\
I_B &= \frac{4T}{2\pi c^2}\int_{-\infty}^{\infty}\omega^2 S_{nn}^2(\omega)d\omega, \qquad (6.10)\\
I_D &= -\frac{1}{2\pi c^2}\int_{-\infty}^{\infty}\omega^2 S_{ss}(\omega)d\omega, \qquad (6.11)\\
q_m &= \frac{\vec{p}_m\cdot\vec{e}}{\parallel\vec{p}_m\parallel}, \qquad (6.12)
\end{aligned}
$$

with $S_{nn}(\omega)$ representing the common noise power spectrum for all the sensors and $S_{ss}(\omega)$ the source signal energy spectrum.

This expression is meaningful in that it not only reflects the performance of the considered estimator, but also provides the means to optimize it. By minimizing (6.5), optimal amplitude weights can be found analytically for the AWEGS algorithm.

6.3.3 Optimal Weight Expression

References [13–18] discuss various optimization approaches in application areas such as time delay estimation, bearing and range estimation, and image registration, etc. Though their specific applications of interest are not the same, their approaches to problem solving follow the same pattern. By defining a proper quantitative measuring index for the application, e.g., the system SNR, the peak width of the system output, or the system error variance, one can improve the performance of an estimation system in terms of these different indices, either numerically as in [8] and [19], or analytically, if possible, as in [13–15]. It is natural that the analytical approach is preferred for realtime applications in practice, since it requires less time and/or resources, compared to numerical approaches.

In this subsection, the SNR approach is studied first as it is intuitive. To decide if the closed-form weight expression obtained is optimal or just suboptimal, the error variance approach is then investigated.

The SNR Approach

When an analytical error variance expression for an estimation algorithm is not available, the SNR is widely used to study the system performance.

With each sensor signal being weighted, a system signal-to-noise ratio denoted $SNR(\vec{w})$ can be defined as a function of weights, and a weight expression that maximizes this $SNR(\vec{w})$ is obtained:

$$\frac{w_m}{w_l} = \frac{d_l}{d_m}, \qquad m, l = 1,\ 2, \cdots M. \tag{6.13}$$

Please refer to Appendix A for detailed derivation.

In [20], the problem of sensor array gain optimization is considered preliminarily using the system SNR. However, near-field and far-field cases are not distinguished in its general model. The difference between various signal processing procedures is not considered, either. The solution given by (4.18) in [20] is hence far from being ready for practical use. As a result, (6.13) can be viewed as complementary to [20].

Since there is no mathematical expression available to link the SNR to the error estimation variance, whether (6.13) is optimal in terms of error variance which directly defines the performance of an estimator remains a question.

The Error Variance Approach

Being the most suitable performance indicator, the error variance will reach its minimum when optimal weights are applied to weight sensor signals. This minimum is defined with the assumption that the weights only take on positive real values. To obtain the closed-form optimal weight expression, the derivative of the error variance expression, given by (6.5), needs to be taken with respect to each w_m and equated to zero. Then the desired solutions might be reached by solving the resulting M nonlinear simultaneous equations, which, however, are not always promised to be solvable. To verify that the extremum obtained is a minimum of $\sigma_\epsilon^2(\vec{w})$, the second derivative of (6.5) with respect to each w_m has to be positive when the solution of w_m is applied.

Differentiating (6.5) directly will result in an extremely complex expression. Therefore, rearrangement of the terms in (6.5) according to the order of a certain w_p is done first, where p takes the value from 1 to M. A closed-form optimal weight expression is then derived and given by

$$w_m = \frac{1}{d_m \sqrt{\sum_{m=1}^{M} \frac{1}{d_m^2}}}, \quad m = 1,\ 2,\ ...,\ M. \tag{6.14}$$

More detailed derivation leading to (6.14) is presented in Appendix B.

Obviously (6.14) is one form of (6.13). It shows that for amplitude weighting using positive real valued weights, maximizing $SNR(\vec{w})$ is equivalent to minimizing $\sigma_\epsilon^2(\vec{w})$ in the presence of noise as assumed here. This may seem intuitive, but due to the lack of a direct mathematical relationship between the SNR and the error variance expression, no firm conclusion can be drawn until the above analysis is performed. The equivalence between the two approaches means the obtained weights can be applied to improve the performance of any estimation/detection system that utilizes the energy in the sum of sensor output signals.

6.3.4 Non-iterative Weighting Technique

With the closed-form optimal weight expression available, how to calculate and apply the weight values effectively is another challenge. (6.14) uses the parameter to be estimated to calculate the optimizing parameter, which is a common contradiction in the field of optimal signal processing. Traditionally this is dealt with iteratively, where the optimizing parameters themselves have to be estimated first, as indicated in [21] and [22]. Here for example, each w_m has to be given the value of "1" to initialize a grid search. A source location estimate is then reached and applied back to (6.14) to calculate weights that will be used in the next iteration. The search will be stopped when some predefined criteria is met after a certain number of iteration are done. Grid search itself is already calculation intensive, and extra iterations simply worsen the situation even further.

In some cases, the initial estimate of the location can be quite inaccurate. For example, the SNR is not high enough, e.g. lower than -30 dB, and/or the grid is coarse, e.g., greater than 100 m, which is common for an initial search across an area of several tens of square kilometers. In these scenarios, the final estimation result can be quite unsatisfactory even after a number of iteration, because the original estimation error can be enlarged during the iterative process. This compromises much of the benefit brought by the proposed amplitude weighting algorithm.

To address these issues, an alternate non-iterative weighting method is proposed, which was first mentioned primitively in [11]. It brings optimal performance with just one iteration for a given grid search since it does not depend on any estimates. The technique is to associate each hypothesized source location with a different set of weights.

The proposed alternate to calculate the weights is thus given by

$$\tilde{w}_m = \frac{1}{\tilde{d}_m \sqrt{\sum_{m=1}^{M} \frac{1}{\tilde{d}_m^2}}}, \qquad \text{m = 1, 2, ..., M,} \tag{6.15}$$

where $\tilde{w}_m$ is the weight assigned to the m^{th} sensor according to the current hypothesized $\tilde{\vec{p}}_s$ whose corresponding total energy is to be calculated. Note that (6.15) is similar to (6.14), except that the two variables $\tilde{w}_m$ and $\tilde{d}_m$ in (6.15) represent the hypothesized instead of the true values.

The total energy can then be calculated using

$$W(\tilde{\vec{p}}_s, \tilde{\vec{w}}) = \int_0^T \left[\sum_{m=1}^{M} \tilde{w}_m r_m(t + \tilde{\tau}_m) \right]^2 dt. \tag{6.16}$$

It is seen that (6.15) promises that the unique set of optimal weights is applied to the grid point coinciding with the true source position. The closer the hypothesized point is to the true source point, the closer the value of (6.15) is to (6.14) . This means (6.15) converges to (6.14) at the true source location. In addition, Appendix C theoretically proves why utilizing (6.15) itself does not cause the peak of (6.4) to move around and that it actually brings unbiased source position estimates. The underlying basis of this technique is that only when correctly weighting and shifting the sensor signals simultaneously, can the purpose of magnifying the stronger signals and reducing the weaker signals be achieved optimally. Other than that, any wrong shifting according to $\tilde{\vec{p}}_s \neq \vec{p}_s$ will cause some sensor signals to be weighted either in the wrong direction, or in the correct direction but not to an optimal extent. Therefore, non-iterative weighting can take the place of iterative weighting to avoid unnecessary calculation for a given grid and the performance becomes more robust. Monte Carlo simulation has verified this, as seen in Figures 6.3, 6.5 and 6.6.

6.4 Performance Evaluation

In this section, the CRB given in [10] and the closed-form estimation error variance expression developed are used to evaluate and compare the original EGS and proposed AWEGS algorithm theoretically using a generated source signal, regardless how the weights are calculated and applied. Monte Carlo simulation is also performed on the EGS, iterative AWEGS and non-iterative AWEGS algorithms to further demonstrate their behaviors using both the generated and a real microseismic signal.

6.4.1 Theoretical and Monte Carlo Evaluation Using Generated Source Signal

In this evaluation, the source location estimation accuracy is observed along with the change of a sensor's position. The evaluation is done for the scenario where there are four sensors configured as $\vec{p}_1$ = (75, 140, 0) m, $\vec{p}_2 = (x_{p2}$, 140, 0) m, $\vec{p}_3$ = (-125, -60, 0) m, $\vec{p}_4$ = (75, -60, 0) m, and the source $\vec{p}_s$ is located at (-25, 40, 0) m. The procedure is to place the second sensor $\vec{p}_2$ at different locations along a horizontal line defined by $y = 140$ m, from the point $x = -1425$ m to the point $x = 75$ m. The configuration is depicted in Figure 6.2.

For any Sensor 2 position, the SNR at Sensor 1 is kept the same and set to be 0 dB. Given a certain error direction $\vec{e}$ = [0.53 0.848 0], which can be arbitrary in practice but is calculated using the source position here, the performance of the EGS and AWEGS is plotted as well as the corresponding CRB in Figure 6.3. This result is generated assuming a flat signal spectrum with a bandwidth of 152.5 Hz, and a signal duration of 0.33 second. The grid search step is 0.16 m.

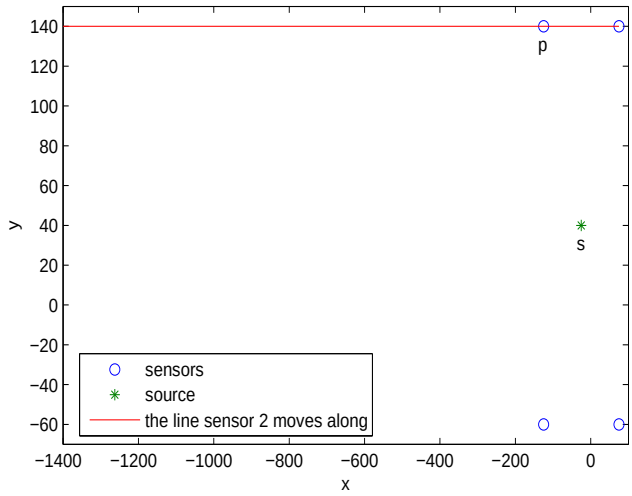

Figure 6.2: Sensor configuration for the simulation using generated source signal

The comparison in Figure 6.3 shows the EGS and AWEGS performance is very similar at the right most 5 points. The overlapping of the root mean square (RMS) error occurs at the point where the X-coordinate of Sensor 2 is -125 m, since the four sensors are symmetrical around the source, meaning that the AWEGS is equivalent to EGS at this point. The AWEGS curve is still pessimistic compared with that of the corresponding CRB, due to the SNR value adopted for the evaluation being only about 0 dB.

It is seen that there is a small gap between the theoretical and Monte Carlo simulation results. This is because the system SNR is only about 0 dB and the time-bandwidth product about 50, which causes some approximation error during the theoretical EEV calculation. The gap is larger for the EGS than the AWEGS, since weighting improves the system SNR, which makes the error analysis more accurate. The gap is also larger when Sensor 2 is further away from the source, since it results in lower system SNR, and hence, poorer theoretical approximation accuracy.

6.4.2 Monte Carlo Evaluation Using Real Signal

This section presents the Monte Carlo simulation performed using more practical parameters. The source signal used for this simulation, as given in Figure 6.4, is a sensor signal recorded in the Lanigan potash mine in Saskatchewan, Canada. The DC component has been removed. Ten sensors that are arbitrarily placed in the Lanigan mine are adopted for the simulation. Their locations are given by (37834, 32438, 0) m, (36200, 33491, 0) m, (35874, 33166, 0) m, (34320, 33834, 0) m, (37136, 33143, 0) m, (34531, 31822, 0) m, (35189, 34891, 0) m, (34434, 32002, 0) m, (35570, 32876, 0) m, and (37386, 34680, 0) m. The sound propagation speed is 2581 m/s. The microseismic event is assumed to take place at the location (26000, 34299, 0) m. The particular direction of the mine tunnel considered in the study, denoted as a unit vector $\vec{e}$, is assumed

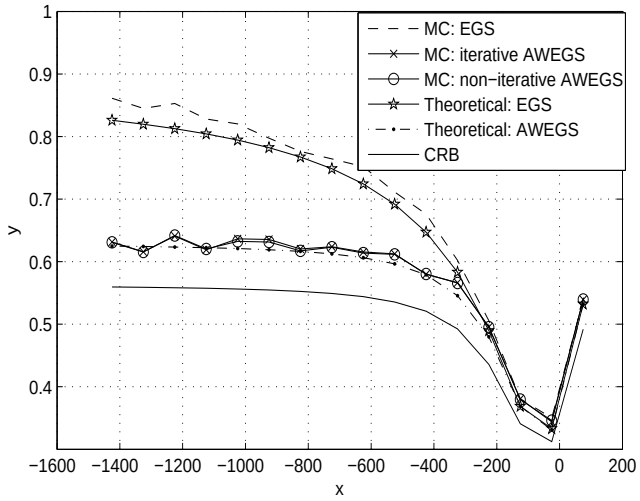

Figure 6.3: Calculated and simulated RMS error for different sensor configurations

to be [0.7240 0.6898 0]. For each SNR value at the nearest sensor, as shown in the plots below, a Monte Carlo simulation of 1000 iteration is performed.

It is seen from Figure 6.5 (taken from [11] for easy reference) that the original EGS algorithm provides good estimation accuracy, with the RMS error being only about 1 m when the SNR is 0 dB. Though the improvement brought by amplitude weighting is large in terms of percentage, the absolute value does not seem to be significant enough to justify its importance. However, it should be noted that the above evaluation is done specifically to demonstrate how close the algorithm performance can reach the CRB in theory. Exact optimal weights that are calculated directly using the true source location and a 10-point interpolation on the source signal are adopted to enable a search over a very fine grid. When the SNR is above 0 dB, the step of the grid is set as small as 0.04 m for the simulation. In practice, this usually takes place only after a smaller area believed to contain the true source position has been identified using coarser grid searches. In addition, the true source position is artificially made to coexist with a grid point in the simulation, which is not always the case in reality. During a coarse grid search, the absolute difference between the unweighted and weighted algorithm in estimation accuracy can be significant. To show this, Monte Carlo simulation is performed for a coarse grid scenario and the result is plotted in Figure 6.6 for non-iterative weighting, iterative weighting with all weights equal to 1 (the first iteration), and iterative weighting using an inaccurate source position estimate that is 51 m away from the true source position (the second iteration). In this simulation, the step of the grid is set at 103 m and the true source position is assumed to be located at the midpoint of two adjacent grid points. In this case, the error variance is much worse than the CRB, and hence the CRB is not plotted together with the simulation results in Figure 6.6.

Non-iterative weighting is demonstrated to outperform the original EGS algorithm by several tens of

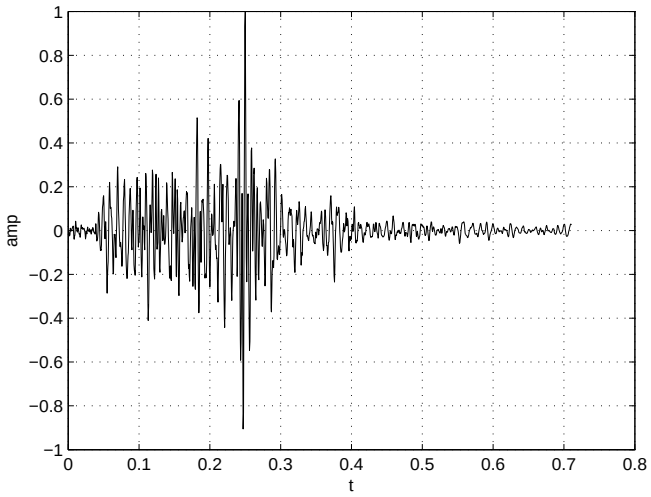

Figure 6.4: Sample sensor signal for the simulation

meters in terms of the RMS. Also, the error in the estimated source position is seen to make iterative weighting much less effective than non-iterative weighting under low SNR's, as expected. It is noticed as well that when the SNR value is greater than a certain value, the RMS error of either non-iterative or iterative weighting becomes zero. The reason is that the grid step of 103 m is too large to reflect the smaller real estimation error under comparatively higher SNR values. Despite that the RMS error is zero, an estimation error of about 51 m exists since the source position is assumed to be at the mid point of two grid points. In practice, the likelihood of the true source position not being overlapped by any grid point is high when a coarse grid is adopted. When an erroneous source estimate, either due to low SNR's or the grid being too coarse, is applied to initialize the following finer grid search, the result can be unsatisfactory.

In summary, the simulation results indicate that both the iterative and non-iterative AWEGS estimators can bring great improvement over the original EGS algorithm. Though in Figure 6.5 the performance plots of the two almost overlap, what is indicated there represents the best possible performance the iterative AWEGS may achieve after a certain number of iteration in practice. Figure 6.6 demonstrates how much difference there can be between them when the grid is coarse. The non-iterative weighting is preferred in the real world because it does not require the true source position to be known in order to perform weighting. This eliminates extra calculation as well as estimation error that can be introduced by iterative weighting.

6.5 Conclusion

This paper focuses on optimal amplitude weighting for source location estimation involving signals of small time-bandwidth product using sparsely placed passive sensors. A closed-form optimal weight expression is

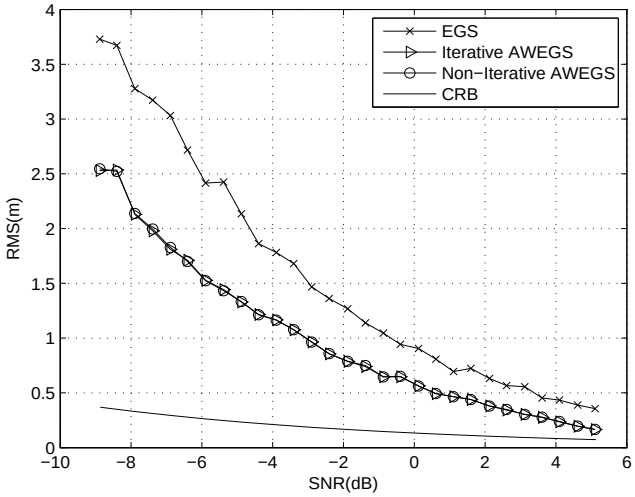

Figure 6.5: Simulated RMS error for a real microseismic event signal (fine grid)

developed and a simple way of weight application is proposed for practical usage.

Based on [8], an analytical error variance expression is presented for the AWEGS algorithm. The traditional functional minimization method involving the concept of multivariate polynomials is used to derive analytically the weight expression by directly minimizing the error variance expression of the AWEGS algorithm.

Based on the proof that there is only one set of optimal weights, the alternate non-iterative amplitude weighting method is proposed to optimize the source estimation accuracy without having to estimate the source position first. Compared with traditional iterative amplitude weighting that is prone to introducing extra error when the SNR is low or the grid is coarse, the alternate one is more attractive for practical implementation. Monte Carlo simulation performed verified this conclusion.

Since an analytical estimation error variance expression is not always available and the SNR is closely linked to system performance, it is intuitive that the weights maximizing the SNR can at least improve the estimation accuracy. This is why the SNR approach was investigated initially. Derivation shows that the SNR and error variance approach are equivalent under the assumed system model. Hence the optimal weights obtained are applicable to improving any algorithms that utilize the energy in the sum of sensor output signals, besides the particular EGS localization algorithm.

The algorithms are evaluated and compared with the corresponding CRB in several scenarios. Both the theoretical evaluation and Monte Carlo simulation results demonstrate that the optimal weights improve the estimation accuracy effectively and can bring it fairly close to the CRB.

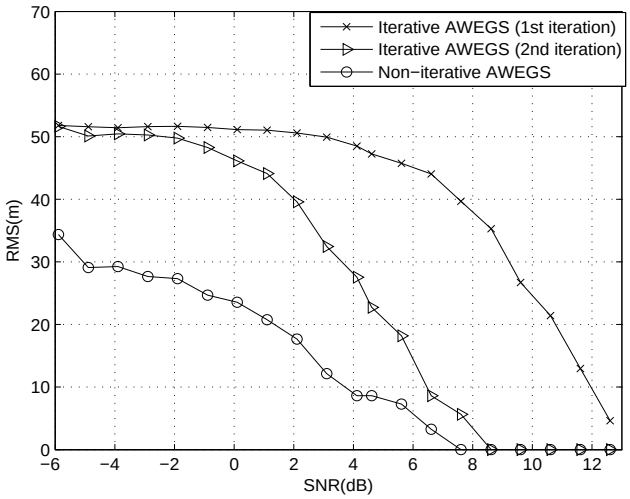

Figure 6.6: Simulated RMS error for a real microseismic event signal (coarse grid)

Similar ideas of using "weighted sum of cross-correlation" and "gain ratios of arrival" are presented in [21] and [22] for different applications using different system models. Compared with them, this work is necessary and unique in that it uncovers the mathematical relationship between the SNR and estimation error variance, in addition to the simple technique of applying the optimal weights.

Acknowledgements

The authors would like to thank TRlabs, Saskatoon, and the Potash Corporation of Saskatchewan Inc. for supporting this research.

A Weight Derivation Using The SNR

Given (6.2) and (6.3), expand (6.4) and take the expected value to get

$$\begin{aligned} W(\tilde{\vec{p}}_s, \vec{w}) &= \int_0^T \left\{ \left[\sum_{m=1}^{M} \frac{w_m}{d_m} s_0(t + \Delta\tilde{\tau}_m) \right]^2 + 2 \sum_{m=1}^{M} \sum_{k=1}^{M} \frac{w_m}{d_m} s_0(t + \Delta\tilde{\tau}_m) w_k \mathbf{n}_k(t + \Delta\tilde{\tau}_k) \right. \\ &+ \left. \left[\sum_{m=1}^{M} w_m \mathbf{n}_m(t + \Delta\tilde{\tau}_m) \right]^2 \right\} dt, \end{aligned} \tag{A.1}$$

$$E[W(\tilde{\vec{p}}_s, \vec{w})] = \int_0^T \left\{ \left[\sum_{m=1}^{M} \frac{w_m}{d_m} s_0(t + \Delta\tilde{\tau}_m) \right]^2 + \left[\sum_{m=1}^{M} E[w_m \mathbf{n}_m(t + \Delta\tilde{\tau}_m)] \right]^2 \right\} dt. \tag{A.2}$$

Shifted versions of the independent noise processes $\mathbf{n}_m(t)$ are used above to make derivation below more compact. Introducing a shift of τ_m does not affect the statistics of the noise and therefore does not change the model.

Since $s_0(t)$ is sensitive to shifting while $n_m(t)$ is not, maximizing the ratio between the signal component and the noise component of (A.2) when $\Delta\tilde{\tau}_m = 0$ is expected to bring a larger decrease in the value of $W(\tilde{\vec{p}}_s, \vec{w})$ when $\Delta\tilde{\tau}_m \neq 0$, and eventually improve the system performance. Hence an SNR is defined as

$$SNR(\vec{w}) = \frac{\int_0^T \left[\sum_{m=1}^{M} \frac{w_m}{d_m} s_0(t) \right]^2 dt}{E\left\{ \int_0^T \left[\sum_{m=1}^{M} w_m \mathbf{n}_m(t) \right]^2 dt \right\}}. \tag{A.3}$$

Under the signal and noise model assumed, $SNR(\vec{w})$ becomes

$$SNR(\vec{w}) = C \cdot \frac{\left(\sum_{m=1}^{M} \frac{w_m}{d_m} \right)^2}{\sum_{m=1}^{M} w_m^2}, \tag{A.4}$$

where

$$C = \frac{\int_0^T s_0^2(t) dt}{\int_0^T E[\mathbf{n}_m^2(t)]}. \tag{A.5}$$

To maximize $SNR(\vec{w})$, $\frac{\partial SNR(\vec{w})}{\partial w_p}$ needs to be equated to zero, which gives

$$\begin{aligned} &\sum_{i=1}^{M} w_i^2 \cdot \frac{2}{d_p} \sum_{i=1}^{M} \frac{w_i}{d_i} - \left(\sum_{i=1}^{M} \frac{w_i}{d_i} \right)^2 \cdot 2w_p = 0, \\ &p = 1,\ 2,\ ...,\ M. \end{aligned} \tag{A.6}$$

followed by

$$w_p = \frac{\frac{1}{dp}\sum_{i=1}^{M} w_i^2}{\sum_{i=1}^{M} \frac{w_i}{d_i}}, \qquad p = 1,\ 2,\ ...,\ M. \tag{A.7}$$

(A.7) represents M simultaneous equations that are not completely independent. Observe that whatever value each w_p is, both $\sum_{i=1}^{M} w_i^2$ and $\sum_{i=1}^{M} \frac{w_i}{d_i}$ are constant for each of the M equations, (A.7) can be simplified as

$$\frac{w_m}{w_l} = \frac{d_l}{d_m}, \qquad m, l = 1,\ 2, \cdots M, \tag{A.8}$$

whose number of linearly independent equations is $M - 1$, instead of M.

B Weight Derivation Using the Error Variance

Due to the complexity of (6.5), terms are rearranged according to the order of w_p ($p = 1, 2, ..., M$). The terms of w_p with orders higher than two all combine to zero. The lengthy procedure of term arranging is omitted to limit the length of the paper. Taking the first derivative of the rearranged (6.5) with respect to each w_p and equating it to zero, the following M simultaneous equations result:

$$\begin{aligned}
&[2C_D^0(I_A C_A^2 + I_B C_B^2) - I_A C_D^1 C_A^1]w_p + I_A C_D^0 C_A^1 \\
&-2C_D^1(I_A C_A^0 + I_B C_B^0) = 0, \\
&p = 1, 2, ..., M,
\end{aligned} \tag{B.1}$$

where

$$\begin{aligned}
C_A^2 \;=\; & \left(\sum_{m=1,m\neq p}^{M} \frac{w_m q_m}{d_m}\right)^2 + \left(\frac{q_p}{d_p}\right)^2 \sum_{k=1,k\neq p}^{M} w_k^2 \\
& + \left(\sum_{m=1,m\neq p}^{M} \frac{w_m q_p}{d_m}\right)^2 + \sum_{k=1,k\neq p}^{M} \left(\frac{w_k q_k}{d_p}\right)^2 \\
& -2 \sum_{l=1,l\neq p}^{M} \frac{w_l}{d_l} \sum_{m=1,m\neq p}^{M} \frac{w_m q_m}{d_m} q_p \\
& -2 \sum_{k=1,k\neq p}^{M} w_k^2 q_k \frac{q_p}{d_p^2},
\end{aligned} \tag{B.2}$$

$$\begin{aligned}
C_A^1 \;=\; & 2\left[\sum_{m=1,m\neq p}^{M} \frac{w_m q_m}{d_m} \sum_{k=1,k\neq p}^{M} w_k^2 \frac{q_p}{d_p}\right. \\
& + \frac{1}{d_p} \sum_{m=1,m\neq p}^{M} \frac{w_m}{d_m} \sum_{k=1,k\neq p}^{M} (w_k q_k)^2 \\
& \left. - \left(\sum_{l=1,l\neq p}^{M} \frac{w_l}{d_l}\frac{q_p}{d_p} + \sum_{l=1,l\neq p}^{M} \frac{w_l q_l}{d_l d_p}\right) \sum_{k=1,k\neq p}^{M} w_k^2 q_k\right],
\end{aligned} \tag{B.3}$$

$$
\begin{aligned}
C_A^0 &= \left(\sum_{m=1,m\neq p}^{M} \frac{w_m q_m}{d_m}\right)^2 \sum_{k=1,k\neq p}^{M} w_k^2 \\
&\quad + \left(\sum_{m=1,m\neq p}^{M} \frac{w_m}{d_m}\right)^2 \sum_{k=1,k\neq p}^{M} (w_k q_k)^2 \\
&\quad -2\left(\sum_{l=1,l\neq p}^{M} \frac{w_l}{d_l} \sum_{m=1,m\neq p}^{M} \frac{w_m q_m}{d_m} \sum_{k=1,k\neq p}^{M} w_k^2 q_k\right),
\end{aligned} \tag{B.4}
$$

$$
\begin{aligned}
C_B^2 &= q_p^2 \sum_{l=1,l\neq p}^{M} w_l^2 + \sum_{m=1,m\neq p}^{M} (w_m q_m)^2 \\
&\quad -2 \sum_{m=1,m\neq p}^{M} w_m^2 q_m q_p,
\end{aligned} \tag{B.5}
$$

$$
\begin{aligned}
C_B^0 &= \sum_{m=1,m\neq p}^{M} (w_m q_m)^2 \sum_{l=1,l\neq p}^{M} w_l^2 \\
&\quad - \left(\sum_{m=1,m\neq p}^{M} w_m^2 q_m\right)^2,
\end{aligned} \tag{B.6}
$$

$$
\begin{aligned}
C_D^1 &= \sum_{m=1,m\neq p}^{M} \frac{w_m q_m^2}{d_m d_p} + \sum_{l=1,l\neq p}^{M} \frac{w_l}{d_l} \frac{q_p^2}{d_p} \\
&\quad -2 \sum_{m=1,m\neq p}^{M} \frac{w_m q_m}{d_m} \frac{q_p}{d_p},
\end{aligned} \tag{B.7}
$$

$$
\begin{aligned}
C_D^0 &= \sum_{m=1,m\neq p}^{M} \frac{w_m q_m^2}{d_m} \sum_{l=1,l\neq p}^{M} \frac{w_l}{d_l} \\
&\quad - \left(\sum_{m=1,m\neq p}^{M} \frac{w_m q_m}{d_m}\right)^2.
\end{aligned} \tag{B.8}
$$

Further assorting (B.1) gives

$$
\begin{aligned}
&[(2C_D^0 C_A^2 - C_D^1 C_A^1) w_p + (C_D^0 C_A^1 - 2C_D^1 C_A^0)] I_A \\
&+[(2C_D^0 C_B^2 w_p) - (2C_D^1 C_B^0)] I_B = 0, \\
&p = 1,\, 2,\, ...,\, M.
\end{aligned} \tag{B.9}
$$

The AWEGS algorithm given by (6.4) implies that the features of specific signal and noise are not relevant. Thus the optimal weights will be applicable to an arbitrary type of source signals, which means (B.9) has to

hold true for any I_A and I_B. As a result, the single equation (B.9) can be broken down into two, (B.10-i) and (B.10-ii):

$$\begin{cases} 2C_D^0 C_B^2 w_p - 2C_D^1 C_B^0 = 0; & \text{(i)} \\ (2C_D^0 C_A^2 - C_D^1 C_A^1) w_p + (C_D^0 C_A^1 - 2C_D^1 C_A^0) = 0. & \text{(ii)} \end{cases}$$
$$p = 1,\, 2,\, ...,\, M. \tag{B.10}$$

If (B.10-i) and (B.10-ii) are linearly independent, an unwanted extra constraint will be introduced and (B.10) will be unsolvable. To avoid this, (B.10-i) and (B.10-ii) must have at least one solution in common.

It is noticed that the left side expressions of (B.10-i) and (B.10-ii) are both multivariate polynomials of w_p with the same degree of 5. For notation simplicity, the two multivariate polynomials are denoted $f(\vec{w})$ and $g(\vec{w})$ for (B.10-i) and (B.10-ii), respectively, in the following context.

One way to make (B.10) solvable is to have

$$\lambda f(\vec{w}) = g(\vec{w}), \tag{B.11}$$

where λ bears an arbitrary real value. (B.11) represents the equality between two polynomials, which by definition is the equality between corresponding terms of the two polynomials. If this is the case, roots of $f(\vec{w})$ will be all and the only roots to $g(\vec{w})$, and the coefficients of all the corresponding terms in the two polynomials are only different by a constant factor of λ.

Start with $f(\vec{w})$ since it looks simpler than $g(\vec{w})$. Since the value of w_p is limited to be real positive, it is expected that it may help solve the problem if every term of $f(\vec{w})$ is positive. Observing C_D^0 given by (B.8) indicates

$$C_D^0 = \sum_{m,l=1;m,l\neq p}^{M} \frac{w_m w_l}{d_m d_l} q_m (q_m - q_l). \tag{B.12}$$

(B.12) seems to be in its simplest form already, but its sign has not been clear to tell yet. Further observing (B.12), a symmetric pattern is revealed among the terms of the double summation and can be used to rearrange (B.12) into

$$C_D^0 = \frac{1}{2} \sum_{m,l=1;m,l\neq p}^{M} \frac{w_m w_l}{d_m d_l} (q_m - q_l)^2, \tag{B.13}$$

which guarantees the positivity. Similarly, C_B^0 given by (B.6) can be rewritten as

$$C_B^0 = \frac{1}{2} \sum_{m,l=1;m,l\neq p}^{M} w_m^2 w_l^2 (q_m - q_l)^2. \tag{B.14}$$

Comparatively easier, C_B^2 and C_D^1 given by (B.5) and (B.7) respectively are sorted into

$$C_B^2 = \sum_{\substack{m\neq p \\ m=1}}^{M} [w_m(q_p - q_m)]^2, \tag{B.15}$$

$$C_D^1 = \sum_{\substack{m\neq p \\ m=1}}^{M} \frac{w_m}{d_m d_p}(q_m - q_p)^2. \tag{B.16}$$

Substitute (B.13)–(B.16) above into (B.10-i) to give

$$\begin{aligned} f(\vec{w}) = & \sum_{m,l,k=1;m,l,k\neq p}^{M}\sum\sum \left(\frac{w_p w_k}{d_m d_l} - \frac{w_m w_l}{d_k d_p}\right) \\ & \times w_m w_l w_k (q_m - q_l)^2 (q_k - q_p)^2, \\ & p = 1,\ 2,\ ...,\ M. \end{aligned} \tag{B.17}$$

Due to the existence of C_A^0 given by (B.4), manipulation on $g(\vec{w})$ shows that $g(\vec{w})$ does not have exactly the same terms as $f(\vec{w})$ does. This indicates that there exists no such λ with a real value to make (B.11) true. The procedure of arranging $g(\vec{w})$ is omitted due to its straightforwardness.

The other way to have (B.10-i) and (B.10-ii) bear a common solution is to force $f(\vec{w})$ and $g(\vec{w})$ to have at least one common zero or equivalently, one common factor. This means both $f(\vec{w})$ and $g(\vec{w})$ have to be capable of being factored. Since in general $q_1 \neq q_2 \neq \cdots \neq q_M$, the factor $\left(\frac{w_p w_k}{d_m d_l} - \frac{w_m w_l}{d_k d_p}\right)$ must be equal for each term contained in the triple summation for (B.17) to be capable of being factored. What are looked for are roots of $f(\vec{w})$, hence this factor has to equal zero. The remaining $w_m w_l w_k (q_m - q_l)^2 (q_k - q_p)^2$ is always positive except when $m = l$, and hence there will be no root coming out of it. Therefore

$$\frac{w_k w_p}{d_m d_l} = \frac{w_m w_l}{d_k d_p}, \qquad k, l, m, p = 1,\ 2,\ ...,\ M. \tag{B.18}$$

(B.18) stands for the unique root of $f(\vec{w})$. It represents $M - 1$ independent simultaneous equations that make the original (B.10-i) true. It implies that only the relative relation between weights matters. The specific value of w_m can be determined by any constraint one wants to put on the weights. Observing $k, l, m, p = 1,\ 2,\ \ldots,\ M$ gives

$$w_l d_l = w_m d_m = a, \qquad\qquad m, l{=}1,\ 2,\ \ldots,\ M, \tag{B.19}$$

where a is a constant with an arbitrary real positive value.

So far, it is revealed that (B.19) represents the only root of $f(\vec{w})$. It is expected to be the root of $g(\vec{w})$ as well, even if it might not be the only one.

Substitute (B.19) into the expression of C_A^0 given by (B.4) to give

$$C_A^0 = a^4 \sum_{k=1,k\neq p}^{M} \frac{1}{d_k^2} \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{q_m(q_m - q_l)}{d_m^2 d_l^2}. \tag{B.20}$$

Attributed to the same symmetric property appeared earlier in (B.13) and (B.14), (B.20) becomes

$$2C_A^0 = a^4 \sum_{k=1,k\neq p}^{M} \frac{1}{d_k^2} \sum\sum_{m,l=1;m,l\neq p}^{M} \left(\frac{q_m - q_l}{d_m d_l}\right)^2, \tag{B.21}$$

after being multiplied by a factor of 2 on both sides.

Substituting (B.19) into C_D^1 given by (B.7) gives

$$C_D^1 = \sum_{\substack{m\neq p \\ m=1}}^{M} \frac{a}{d_m^2 d_p}(q_m - q_p)^2. \tag{B.22}$$

Therefore,

$$\begin{aligned} & -2C_D^1 C_A^0 \\ = \; & -\frac{a^5}{d_p} \sum_{k=1,k\neq p}^{M} \left(\frac{q_k - q_p}{d_k}\right)^2 \sum_{n=1,n\neq p}^{M} \frac{1}{d_n^2} \\ & \times \sum\sum_{m,l=1;m,l\neq p}^{M} \left(\frac{q_m - q_l}{d_m d_l}\right)^2. \end{aligned} \tag{B.23}$$

Similarly, (B.13) and (B.3) can be resorted to give the product

$$\begin{aligned} & C_D^0 C_A^1 \\ = \; & \frac{a^5}{d_p} \sum_{k=1,k\neq p}^{M} \frac{(q_p - q_k)(q_n - q_k)}{d_k^2 d_n^2} \sum\sum_{m,l=1;m,l\neq p}^{M} \left(\frac{q_m - q_l}{d_m d_l}\right)^2. \end{aligned} \tag{B.24}$$

Hence

$$\begin{aligned} & C_D^0 C_A^1 - 2C_D^1 C_A^0 \\ = \; & -\frac{a^5}{d_p} \sum\sum_{m,l=1;m,l\neq p}^{M} \left(\frac{q_m - q_l}{d_m d_l}\right)^2 \left(\sum_{\substack{k\neq p \\ k=1}}^{M} \frac{q_p - q_k}{d_k^2}\right)^2. \end{aligned} \tag{B.25}$$

The following step is to show that $(2C_D^0 C_A^2 - C_D^1 C_A^1)w_p$ has the same absolute value as that given by (B.25), but with a positive sign. First look at C_A^1 given by (B.3). Taking advantage of the factor 2, C_A^1 can be rewritten in the form of a sum consisting of two identical parts, as indicated in the two square brackets

below, where index m in one part corresponds to index l in the other part,

$$\begin{aligned} C_A^1 &= \Bigg[\sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_m w_l^2 q_m q_p}{d_m d_p} + \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_m w_l^2 q_l^2}{d_m d_p} \\ &\quad - \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_m w_l^2 q_l q_p}{d_m d_p} - \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_m w_l^2 q_m q_l}{d_m d_p} \Bigg] \\ &\quad + \Bigg[\sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_l w_m^2 q_l q_p}{d_l d_p} + \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_l w_m^2 q_m^2}{d_l d_p} \\ &\quad - \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_l w_m^2 q_m q_p}{d_l d_p} - \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_l w_m^2 q_l q_m}{d_l d_p} \Bigg]. \end{aligned} \tag{B.26}$$

(B.26) can be further arranged into a more compact form

$$\begin{aligned} C_A^1 &= \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_m w_l q_p}{d_m d_l d_p}(d_l w_l - d_m w_m)(q_m - q_l) \\ &\quad + \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{w_m w_l}{d_m d_p d_l}(d_m w_m q_m - d_l w_l q_l)(q_m - q_l). \end{aligned} \tag{B.27}$$

(B.19) makes the first double summation term in (B.27) equal to zero and hence (B.27) is reduced to

$$C_A^1 = \sum_{m,l=1;m,l\neq p}^{M}\sum \frac{a^3 (q_m - q_l)^2}{d_m^2 d_l^2 d_p}. \tag{B.28}$$

Next observe C_A^2 given by (B.2), which can also be arranged into a more compact form made up of two positive parts,

$$C_A^2 = \left[\sum_{\substack{m=1 \\ m\neq p}}^{M} \frac{w_m}{d_m}(q_m - q_p) \right]^2 + \sum_{\substack{m=1 \\ m\neq p}}^{M} \left[\frac{w_m}{d_p}(q_p - q_m) \right]^2. \tag{B.29}$$

Now put (B.13), (B.16), (B.28), and (B.29) together and substitute with (B.19) to give

$$\begin{aligned} &(2C_D^0 C_A^2 - C_D^1 C_A^1) w_p \\ &= \frac{a^5}{d_p} \sum_{m,l=1;m,l\neq p}^{M}\sum \left(\frac{q_m - q_l}{d_m d_l} \right)^2 \left(\sum_{\substack{k=1 \\ k\neq p}}^{M} \frac{q_p - q_k}{d_k^2} \right)^2, \end{aligned} \tag{B.30}$$

which combines together with (B.25) to yield a zero.

By now it has been proven that (B.19) is also the solution to (B.10-ii). Hence (B.19) brings the only extremum of (6.5). However, it is not certain yet if this extremum is the desired minimum or the undesired

maximum of $\sigma_\epsilon^2(\vec{w})$. To complete this minimization problem mathematically, next it is proven that the slope sign of the line determined by (B.1), as given by the second derivative of the rearranged (6.5), is always positive when (B.19) is applied.

Since I_A, I_B, C_B^2, and $C_{A\_1}^2$ (the first square term in C_A^2 given by (B.29)) are all greater than 0, there is

$$\begin{aligned} 2C_D^0(I_A C_A^2 + I_B C_B^2) &> 2C_D^0 I_A C_A^2 \\ &> 2I_A C_D^0 C_{A\_2}^2, \end{aligned} \tag{B.31}$$

where $C_{A\_2}^2$ denotes the second square term in (B.29). It is found that $2I_A C_D^0 C_{A\_1}^2 - I_A C_D^1 C_A^1 = 0$ by utilizing (B.19), and hence

$$\begin{aligned} \frac{\partial^2 \sigma_\epsilon^2(\vec{w})}{\partial w_p^2} &= 2C_D^0(I_A C_A^2 + I_B C_B^2) - I_A C_D^1 C_A^1 \\ &> 2I_A C_D^0 C_{A\_2}^2 \\ &> 0, \end{aligned}$$

$$p = 1,\, 2,\, ...,\, M. \tag{B.32}$$

The inequality given by (B.32) proves that the solution given by (B.19) will bring the only global minimum of $\sigma_\epsilon^2(\vec{w})$.

Applying the constraint of $\sum_{m=1}^{M} w_m^2 = 1$ leads to (6.14). It keeps the noise component of $W(\tilde{\vec{p}}_s, \vec{w})$ statistically unchanged for each grid point, as indicated in Appendix C.

C Non-iterative Weighting

Similar to (9)–(11) of [8], there is

$$W(\tilde{\vec{p}}_s, \tilde{\vec{w}}) = W_s(\tilde{\vec{p}}_s, \tilde{\vec{w}}) + W_n(\tilde{\vec{p}}_s, \tilde{\vec{w}}), \tag{C.1}$$

where $W_s(\tilde{\vec{p}}_s, \tilde{\vec{w}})$ and $W_n(\tilde{\vec{p}}_s, \tilde{\vec{w}})$ are the signal and noise components, respectively, given by

$$W_s(\tilde{\vec{p}}_s, \tilde{\vec{w}}) = \sum_{m=1}^{M}\sum_{k=1}^{M}\int_0^T \frac{\tilde{w}_m\tilde{w}_k}{d_m d_k} s_0(t+\Delta\tilde{\tau}_m)s_0(t+\Delta\tilde{\tau}_k)dt, \tag{C.2}$$

$$\begin{aligned} W_n(\tilde{\vec{p}}_s, \tilde{\vec{w}}) = \sum_{m=1}^{M}\sum_{k=1}^{M}\int_0^T \tilde{w}_m\tilde{w}_k \Big[& \frac{1}{d_m} s_0(t+\Delta\tilde{\tau}_m)\mathbf{n}_k(t+\Delta\tilde{\tau}_k) \\ & + \mathbf{n}_m(t+\Delta\tilde{\tau}_m)\frac{1}{d_k}s_0(t+\Delta\tilde{\tau}_k) \\ & + \mathbf{n}_m(t+\Delta\tilde{\tau}_m)\mathbf{n}_k(t+\Delta\tilde{\tau}_k)\Big]\, dt. \end{aligned} \tag{C.3}$$

Since $\tilde{w}_m = \frac{1}{\tilde{d}_m}$ is simple and satisfies (B.19), it is used to analyze how the value of $\tilde{w}_m$ changes when the value of $\tilde{\epsilon}$ changes. According to (5) in [8],

$$\tilde{d}_m \approx d_m - q_m\tilde{\epsilon}, \tag{C.4}$$

there is

$$\frac{d\tilde{w}_m}{d\tilde{\epsilon}} = \frac{-q_m}{(d_m - q_m\tilde{\epsilon})^2}. \tag{C.5}$$

When $|\tilde{\epsilon}\vec{e}| \ll d_m$, $\frac{d\tilde{w}_m}{d\tilde{\epsilon}}$ is very small and negligible. Therefore, based on (14) of [8], the signal component given by (C.2) can be approximated for small $\tilde{\epsilon}$ for any $\vec{e}$ by the second order Maclaurin polynomial

$$W_s(\tilde{\vec{p}}_s, \vec{w}) \approx c_0 + c_2\tilde{\epsilon}^2, \tag{C.6}$$

where

$$
\begin{aligned}
c_0 &= W_s(\tilde{\vec{p}}_s, \vec{w})\Big|_{\tilde{\epsilon}=0} \qquad \text{(C.7)}\\
&= \left(\sum_{m=1}^{M} \frac{w_m}{d_m}\right)^2 \int_0^T s_0^2(t)dt,\\
c_2 &= \left.\frac{d^2 W_s(\tilde{\vec{p}}_s, \vec{w})}{d\tilde{\epsilon}^2}\right|_{\tilde{\epsilon}=0} \qquad \text{(C.8)}\\
&= -\frac{1}{2\pi c^2}\int_{-\infty}^{\infty} \omega^2 S_{ss}(\omega)d\omega \sum_{m=1}^{M}\sum_{k=1}^{M} \frac{w_m w_k}{d_m d_k}\left(\frac{\vec{p}_m\cdot\vec{e}}{|\vec{p}_m|} - \frac{\vec{p}_k\cdot\vec{e}}{|\vec{p}_k|}\right)^2\\
&< 0.
\end{aligned}
$$

Since (6.15) becomes (6.14) at the true source position where $\tilde{\epsilon}=0$, c_0 is the same with either (6.15) or (6.14) applied. Actually using Lagrange multiplier, $\left(\sum_{m=1}^{M}\frac{w_m}{d_m}\right)^2$ in c_0 is only maximal at the true source position with the constraint of $\sum_{m=1}^{M} w_m^2 = 1$ applied. Also $c_2\tilde{\epsilon}^2$ is always non-positive for any grid point, the value of $W_s(\tilde{\vec{p}}_s, \vec{w})$ at $\tilde{\epsilon}=0$ is always greater than that at hypothesized points where $\tilde{\epsilon}\neq 0$, whether the value of w_m is calculated from (6.15) or (6.14).

Taking the expected value of the noise component given by (C.3) leads to

$$
\begin{aligned}
E[W_n(\tilde{\vec{p}}_s, \vec{w})] &= E\left[\sum_{m=1}^{M} \int_0^T w_m^2 \mathbf{n}_m^2(t+\Delta\tilde{\tau}_m)\right]dt \qquad \text{(C.9)}\\
&= \sum_{m=1}^{M} w_m^2 \int_0^T E[\mathbf{n}_m^2(t)]\\
&= \int_0^T R_{nn}(0),
\end{aligned}
$$

where $E[\mathbf{n}_m^2(t)] = R_{nn}(0)$ for $m = 1,\ 2,\ ...,\ M$. Obviously the value of $E[W_n(\tilde{\vec{p}}_s, \vec{w})]$ remains unchanged for any grid point, whether w_m is calculated from (6.15) or (6.14), since both of them satisfy $\sum_{m=1}^{M} w_m^2 = 1$.

So far it is proven statistically that $W(\tilde{\vec{p}}_s, \vec{w})$ will always peak at the true source position, with either (6.15) or (6.14) applied. Both non-iterative weighting and iterative weighting bring unbiased source position estimates. Given the results in Appendix B, it is known that both of them bring minimal error variance in the end.

References

[1] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 24, pp. 320–327, August 1976.

[2] G. C. Carter, *Coherence and Time Delay Estimation.* IEEE Press, 1993.

[3] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 35, pp. 1223–1225, August 1987.

[4] J. O. Smith and J. S. Abel, "The spherical interpolation method of source localization," *IEEE J. of Oceani. Eng.*, vol. 12, pp. 246–252, January 1987.

[5] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. on Signal Processing*, vol. 42, pp. 1905–1915, August 1994.

[6] Y. Huang, J. Benesty, G. Elko, and R. Mersereau, "Real-time passive source localization: a practical linear-correction least-squares approach," *IEEE Trans. on Speech and Audio Processing*, vol. 9, pp. 943–956, November 2001.

[7] J. P. Ianniello, "Large and small error performance limits for multipath time delay estimation," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, pp. 245–251, April 1986.

[8] B. L. Daku and E. J. Salt, "Error analysis of a localization algorithm for finite-duration events," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1024–1034, March 2007.

[9] S. Li and B. Daku, "Performance comparison for three source localization algorithms," in *Proceedings of CCECE-2005*, CCECE at Saskatoon Inn and Conference Centre, Saskatoon, Canada, May 2005.

[10] S. Li and B. L. Daku, "Cramér-rao bound on passive source localization for general gaussian noise," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E93-A, pp. 914–925, May 2010.

[11] S. Li and B. Daku, "Alternate amplitude weighting approach for passive source localization using the energy-based grid search algorithm," in *Proceedings of CCECE-2008*, CCECE at Sheraton Fallsview, Ontario, Canada, May 2008.

[12] J. G. Proakis, *Digital Communications.* McGraw-Hill, 2001.

[13] W. R. Hahn, "Optimum processing for passive sonar range and bearing estimation," *J. Acoust. Soc. Am.*, vol. 58, pp. 201–207, July 1975.

[14] C. D. Mcgillem and M. Svedlow, "Optimum filter for minimization of image registration error variance," *IEEE Trans. on Geoscience Electronics*, vol. GE-15, pp. 257–259, October 1977.

[15] T. L. Steding and F. W. Smith, "Optimum filters for image registration," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-15, pp. 849–860, November 1979.

[16] D. Korompis, K. Yao, and F. Lorenzelli, "Broadband maximum energy array with user imposed spatial and frequency constraints," in *Proceedings of ICASSP-94*, vol. iv, pp. IV529–IV532, ICASSP-94, April 1994.

[17] T. Tung, K. Yao, D. Chen, R. Hudson, and C. Reed, "Source localization and spatial filtering using wideband MUSIC and maximum power beamforming for multimedia applications," in *1999 IEEE Workshop on Signal Processing Systems: Design and Implementation*, (Taipei, Taiwan), pp. 625–634, SIPS 99, 1999.

[18] E. F. Sagiroglu, "Localization of wide-band signals via extended Kalman filter," in *Proceedings of IEEE International Symposium on Circuits and Systems*, vol. 4, pp. IV163–IV166, 1999.

[19] J. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.

[20] D. H. Johnson and D. E. Dudgeon, *Array Signal Processing*. Englewood Cliffs, NJ: Prentice-Hall, 1993.

[21] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. on Signal Processing*, vol. 50, pp. 1843–1854, August 2002.

[22] K. C. Ho and M. Sun, "Source localization using time differences of arrival and gain ratios of arrival," *IEEE Trans. on Signal Processing*, vol. 56, pp. 464–477, February 2008.

7 Frequency Weighting to Minimize Source Location Estimation Error Variance for Short Duration Signals

The original manuscript was conditionally accepted by IEEE Trans. on Signal Processing in Dec. 2013. The revised manuscript included here is to be submitted again to IEEE Trans. on Signal Processing.

In Chapter 6, an amplitude weighting strategy, using a set of real positive values, was proposed to improve the EGS algorithm performance by increasing the passive source localization accuracy. In this chapter, a frequency weighting strategy for the same purpose is proposed and realized in the frequency domain using a filter. The source signal and noise are both assumed stationary in this chapter to ensure that the Fourier transform is applicable to the mathematical derivation. Theoretical derivation of several classical filters given in the literature are also presented. Monte Carlo simulation is performed to compare the performance of all the filters investigated.

Frequency Weighting to Minimize Source Location Estimation Error Variance for Short Duration Signals

Sha Li* and Brian L.F. Daku, *Member, IEEE*

Abstract

In this paper, frequency domain optimization techniques are investigated for passive acoustic source localization to yield more accurate source location estimates. An analytic frequency weighted estimation error variance (EEV) expression is derived using the spectrum representation and stochastic process theory. With this EEV expression, an integral equation is established and solved using the calculus of variations technique. The solution to the equation gives the expression of a filter transfer function that minimizes this frequency domain EEV derived. Ignoring various parts of the EEV expression, during the minimization procedure using Cauchy-Schwarz inequality, leads to several well known classical filter transfer functions that have been developed to deal with source localization problems in the literature. This paper demonstrates that unlike the case of amplitude weighting proposed in one of the authors' earlier papers, the SNR approach and the EEV approach are not equivalent when frequency weighting is used to minimize the EEV. Monte Carlo simulation is performed to evaluate all the filters of interest in this work. The results confirm that the filter obtained by minimizing the analytic EEV expression derived in this paper outperforms all the other filters considered.

Index Terms

acoustic source localization, location estimation, grid search, microseismic.

EDICS Category: SEN-LOCL

The authors are with the Department of Electrical Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: sha.li@usask.ca; brian.daku@usask.ca).

7.1 Introduction

The task of passively locating an acoustic source is required in a variety of applications, such as the tracking of underwater vehicles, video conferencing, mobile robot navigation and seismic event localization. Traditional two-step source location estimation techniques require a time difference of arrival (TDOA) estimation procedure, since they calculate the source location using these TDOA estimates. For example, the well-known technique utilizing the classic generalized cross-correlation (GCC) function to estimate the TDOA is discussed in [1]. A number of suboptimal closed-form techniques have been developed using spherical intersection [2], spherical interpolation [3] and numerical simplex algorithm [4] to identify the source location using the TDOA estimates. For these two-step algorithms to operate above threshold, the received signals need to have sufficient duration. However, for a source that emanates a short duration signal with a small time-bandwidth product, such as a microseismic event, these two-step algorithms cannot consistently guarantee desired performance. Thus alternative approaches must be considered for applications involving this type of source signal. One such approach is the Energy-based Grid Search (EGS) algorithm discussed in [5] that avoids TDOA estimation. The work presented in this paper further improves the estimation accuracy of the original EGS algorithm.

Optimization techniques can be developed in both the time and frequency domain based on various performance indices. The most direct index is the estimation error variance (EEV), which in many cases is difficult to obtain. However, an explicit EEV expression for the EGS algorithm has been given in [5], which was also used to develop an amplitude weighted EGS (AWEGS) algorithm, proposed in [6]. In [6], a closed-form expression is derived to yield real positive weight values that minimizes this EEV expression in the context of amplitude weighting. While the amplitude weighting works regardless of the source signal or additive noise components, a major prerequisite for frequency domain optimization techniques to be practical is that the spectrum information for both the source signal and noise are known or can be estimated with acceptable accuracy.

During the last half a century, many optimization techniques have been developed, most of which are rooted firmly on fundamental frequency domain concepts. An optimal digital filter design based on Capon's minimum variance principle is proposed in [7] to do fundamental frequency estimation. Another optimal filter design based on the Kalman-Bucy filter is proposed in [8] for linear systems with state delay over linear observations. Neither of the filters proposed in [7] and [8] are applicable to the EGS algorithm directly, due to the different parameters that are estimated.

A linearly spaced broadband-sensor array is employed in [9] and [10], to provide "rough maximum energy concentration over some desired spatial 'look' and frequency regions, with user imposed spatial and frequency attenuation constraints". This is a very simple and effective frequency weighting method, but it is proposed for far-field applications, where the direction of arrival (DOA) and the signal strength across the sensor array

is deemed identical. Since the microseismic events considered here take place in comparatively close vicinity to the sensors, the difference in source signal attenuation across the sensor array cannot be ignored, and the DOA of the source signal at each sensor should not be taken as the same.

In many cases, the spectrum information for both the source signal and noise can be estimated quite precisely by observing past events. Under such circumstances, roughly selecting the energy-dominant frequency regions of the signal as described in the previous paragraph is not sophisticated enough to take full advantage of all the information available.

There have been a few optimal signal processing techniques developed in the literature for TDOA/DOA estimation using various kinds of filters. Ideally, with optimized TDOA estimates, optimal passive range and bearing estimation is expected. However, the non-linear relationship between the TDOAs and the source position, and the lack of accommodation for the amplitude differences across the sensors, imply that the filters producing optimal TDOAs for far-field applications may not be the best choice for the application here. Still, it is of interest to see by how much these filters may possibly improve the location estimation accuracy, if they can be applied as the common filter to the EGS algorithm described in Figure 7.2.

One means of optimal TDOA estimation is to pre-filter signals before computing their generalized cross-correlation. When there is no *a priori* information about the source signal and noise spectrum, the most convenient way to sharpen the cross-correlation peak is to whiten the received signals. A deterministic approach is presented in [1] and leads to the so-called phase-correlation technique [11]. A statistical approach to the whitening problem has been discussed in [12], though the application of interest is image registration. Another alternative way to optimize the time delay estimation is to minimize the maximum likelihood (ML) estimate of the TDOA with the aid of a separate correlator delay measurement system [13]. A representative TDOA optimizing filter from [13] is evaluated in this paper.

There has been some work done to study ambient seismic noise to date, e.g. [14] and [15]. However, developing an analytic filter expression for optimal signal processing for the problem of near-field source localization involving seismic signals has received little attention compared to the far-field source application.

The application of registering a received image (with noise) with a known ideal image is investigated in [16–18]. The authors analyzed image registration errors according to different performance measures, such as the overlay quality, the peak-to-sidelobe ratio of the cross-correlation function, and the covariance of the registration error that is influenced by the presence of both pure noise and signal distortion. Based on these measures, the authors derived several filter transfer functions for optimal image registration. In [19], optimal filters are derived for signal detection purpose based on three performance measures, namely: the SNR, localization, and suppression of multiple peaks.

Though, the results presented for the applications mentioned in the paragraph above are different, the

general procedure used in this work is the same: determine a target measure to be optimized, develop its analytic expression, and then optimize this expression that results in a desired filter. As a complementary work to [6], this work follows the same approach to obtain such a filter. It focuses on optimal signal weighting in the frequency domain that can be applied to the EGS algorithm. The major contribution of this paper includes the proposition of a frequency weighted EGS (FWEGS) algorithm for the application of interest, development of an analytic expression that closely approximates the actual FWEGS EEV, and derivation of a filter transfer function optimal in terms of minimizing this closed-form EEV expression given the adopted system model.

The FWEGS algorithm can be applied to a variety of acoustic source localization applications where known or estimated source signals and/or noise exist, such as radar and sonar. In this paper, the FWEGS algorithm performance is evaluated when the following classical filters of interest are applied. Some of the filters below can be obtained using the FWEGS EEV expression arrived in this paper.

Matched Filter. The matched filter has been discussed in great detail in most communications textbooks for systems where the signal to process is a copy of a known signal. The classic matched filter gives the maximum signal-to-noise ratio at a particular time t_0 when the noise is additive white Gaussian. Using spectrum representation, its transfer function is the complex conjugate of the Fourier transform of the known signal with a delay coefficient t_0 constant for all the frequency components.

Noise Whitening Filter. When the transmitted source signal is corrupted with additive wide-sense-stationary coloured Gaussian noise at a receiver, a filter can be used to whiten the noise if the power spectrum of the noise is known. It helps optimize the system performance even when the source signal spectrum is unknown.

Eckart Filter. The Eckart filter is named after the author of [20]. It "maximizes the deflection criterion, i.e., the ratio of the change in mean correlator output due to signal present to the standard deviation of correlator output due to noise alone" [1]. When the noise is white Gaussian or the SNR is very high, the Eckart filter becomes the matched filter. When the SNR is very low, the Eckart filter becomes the noise whitening filter.

Filter for suboptimal Time Delay Estimation. A suboptimal filter expression for time delay estimation with an example involving two sensors is presented in [13]. The expression is modified to accommodate more than two sensors, so it can be applicable under the given system model of this paper and evaluated against other filters of interest. It takes both signal and noise spectrum into consideration, as does the filter minimizing the EEV expression, but is different in terms of mathematical details.

The outline of the paper is as follows: Section 7.2 briefly describes the system model adopted for this study. Section 7.3 discusses how to apply the filter(s). Section 7.4 presents how the analytic expression of the

selected performance measure, the EEV, has been developed. Section 7.5 provides the derivation procedure for several filter transfer functions using this EEV expression and the SNR expression. Then a filter transfer function is derived in Section 7.6 by minimizing the same EEV expression developed in Section 7.4. The performance evaluation for all the filters considered in this work is presented in Section 7.7. A conclusion is provided in Section 7.8.

7.2 System Model

The signal considered in this paper is similar to that in [5], which is a very short duration acoustic wave that is set in motion by a sudden shift or rupture of rock and may only last a fraction of a second. This means multi-path signals received at the sensors within the vicinity of the event are unlikely to overlap, and the signals arriving along indirect paths can be easily identified. The indirect-path signals caused by reflection are much weaker than the direct-path signal and hence the system model adopted in this paper considers only the direct-path scenario.

The possibility of more than one microseismic event taking place side by side at the same time is extremely low, and the radial distance to the sensors is much larger than the dimension of the vibration source generating the acoustic waves. Throughout this paper it is assumed that within a given short period of time, only one microseismic event takes place, which is reasonable considering the length of a microseismic signal. Multiple-source localization is therefore not of concern within the scope of this research.

It is thus assumed that the microseismic signal originates from a single-point; the propagation medium is homogeneous and elastic, where there is no energy loss from friction so the signal energy attenuates proportionally to the square of distance; all the indirect-path signals are ignored; the response of every sensor is identical; the output of each sensor contains additive independent noise; and the noises at all sensors are stationary Gaussian stochastic processes with an identical power spectra.

The geometry of the sensor system is shown in Figure 7.1. The source position is given by $\vec{p}_s = (x_s, y_s, z_s)$. There are M sensors that are placed arbitrarily and sparsely in the system. The position of sensor m is denoted $\vec{p}_m = (x_m, y_m, z_m)$. The distance from the source to sensor m is denoted

$$\begin{aligned} d_m &= \| \vec{p}_m - \vec{p}_s \| \\ &= \sqrt{(x_m - x_s)^2 + (y_m - y_s)^2 + (z_m - z_s)^2}. \end{aligned} \tag{7.1}$$

The output of sensor m, denoted $r_m(t)$, is a continuous sample function of the stochastic process $\mathbf{r}_m(t)$, which is given by

$$\mathbf{r}_m(t) = s_m(t) + \mathbf{n}_m(t), \tag{7.2}$$

where $\mathbf{n}_m(t)$ is an additive random noise process and $s_m(t)$ is the signal received at sensor m. $\mathbf{n}_m(t)$ is

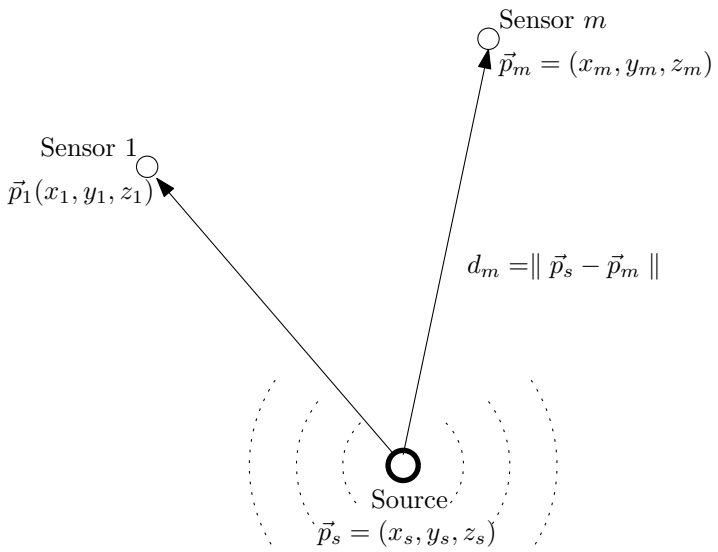

Figure 7.1: Coordinate system for the sensors

Gaussian noise with zero-mean and unknown covariance. Since the sensors are sparsely positioned in the application of interest, the noise can be taken as spatially uncorrelated. $s_m(t)$ is simply a delayed and attenuated version of the original unknown deterministic source signal $s_0(t)$, which is independent of $\mathbf{n}_m(t)$.

It is assumed that all the received signals are collected by sensors that are at least one unit distance away from the source so that in mathematical calculation the received sensor signal is not amplified. The unit can be arbitrary so long as it is in accordance with the distance measure system adopted for the mathematic derivation procedure. Now define

$$s_m(t) \quad = \quad \frac{s_0(t-\tau_m)}{d_m}, \tag{7.3}$$

where $\tau_m = \frac{d_m}{c}$ is the time for the unknown signal $s_0(t)$ to travel from the source location to sensor m and c is the propagation speed of the mechanical wave.

It is also assumed that the real source signal $s_0(t)$ contains no discontinuity, is finite, lasts for a time duration T, and has a bandwidth Ω. Its energy spectrum is denoted $S_{ss}(\omega)$. The start time of the source event is taken as the time origin, and T_o is the end of the observation time, which is assumed to be much greater than T. T_o may be extended to infinity to ensure that no statistical information of the random process is lost. $n_m(t)$ and $r_m(t)$ are both real functions and are used to denote sample functions of $\mathbf{n}_m(t)$ and $\mathbf{r}_m(t)$, respectively. The noise at all sensors has a common auto-correlation function $R_{nn}(\tau)$ and power spectrum $S_{nn}(\omega)$. In addition, for noise processes at sensor l and m, there is

$$E[n_l(t)n_m(t)] = 0, \qquad l \neq m, \tag{7.4}$$

where "$E[\ ]$" represents the expected value.

7.3 Proposed Frequency Weighted EGS (FWEGS) Algorithm

This section discusses possible optimization techniques by taking into account the frequency components of the source signal and the noise. The amplitude weighted EGS algorithm [6] is given by

$$\mathbf{W}(\tilde{\vec{p}}_s) = \int_0^T \left[\sum_{m=1}^{M} w_m r_m(t + \tilde{\tau}_m) \right]^2 dt, \tag{7.5}$$

where $\vec{w} = [w_1\ w_2\ ...\ w_M]$, w_m is the weight for m_{th} sensor signal, $\tilde{\vec{p}}_s$ is the hypothesized source position, and $\tilde{\tau}_m$ is the hypothesized delay for the signal to travel from $\tilde{\vec{p}}_s$ to the m_{th} sensor. Uniform amplitude weight values are assumed within the scope of this paper to focus on frequency weighting, thus $w_m = 1$ for $m = 1$, ..., M.

According to the system model presented in Section 7.2, the noises at all the sensors are assumed independent Gaussian noise processes with the same distribution. In this case, the FWEGS can be implemented as depicted in Figure 7.2, where only one common filter is applied. First, the input sensor signal $s_m(t)$ is mapped into the frequency domain using a Fourier Transform (FT), weighted and delayed by $\tilde{\tau}_m$ that is consistent with the hypothesized source position $\tilde{\vec{p}}_s$. Then, the weighted and shifted sensor signals are summed before the common filter, denoted $F(\omega)$, is applied. This is followed by the squaring and integration steps. The final output is the frequency weighted total energy as a function of the hypothesized source position $\tilde{\vec{p}}_s$. $F(\omega)$ is the Fourier transform of the impulse response of the common filter denoted $f(t)$. Note that an increase in the number of sensors does not affect the way the filter is designed and applied, since the noise at the sensors is assumed to be independent with the same stochastic characteristics.

The FWEGS algorithm proposed, as shown in Figure 7.2, is given by

$$\mathbf{W}(\tilde{\vec{p}}_s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| F(\omega) \sum_{m=1}^{M} w_m \left[e^{j\omega(\tilde{\tau}_m - \tau_m)} \frac{X_s(\omega)}{d_m} + e^{j\omega\tilde{\tau}_m} X_{n_m}(\omega) \right] \right|^2 d\omega, \tag{7.6}$$

where $X_s(\omega)$, $X_{n_m}(\omega)$, and $F(\omega)$ are the Fourier transforms of $s_0(t)$, $n_m(t)$ and $f(t)$, respectively. It is seen that the FWEGS algorithm proposed here is simply the EGS algorithm with a filter applied.

7.4 Derivation of the Performance Measure

As mentioned earlier, various performance measures can be used to optimize an algorithm. The two most widely used measures are the system SNR and the EEV, with the EEV being the most direct performance measure for a parameter estimator. Hence EEV is selected as the target performance measure to optimize in this work. A closed-form EEV expression for the EGS algorithm is presented in [5]. An analytic EEV expression for the FWEGS algorithm as indicated in Figure 7.2 is presented using spectrum representation in (7.7), whose detailed derivation is provided in Appendix D.

$$\sigma_\epsilon^2(F) = \frac{A(F) + B(F)}{D^2(F)}, \tag{7.7}$$

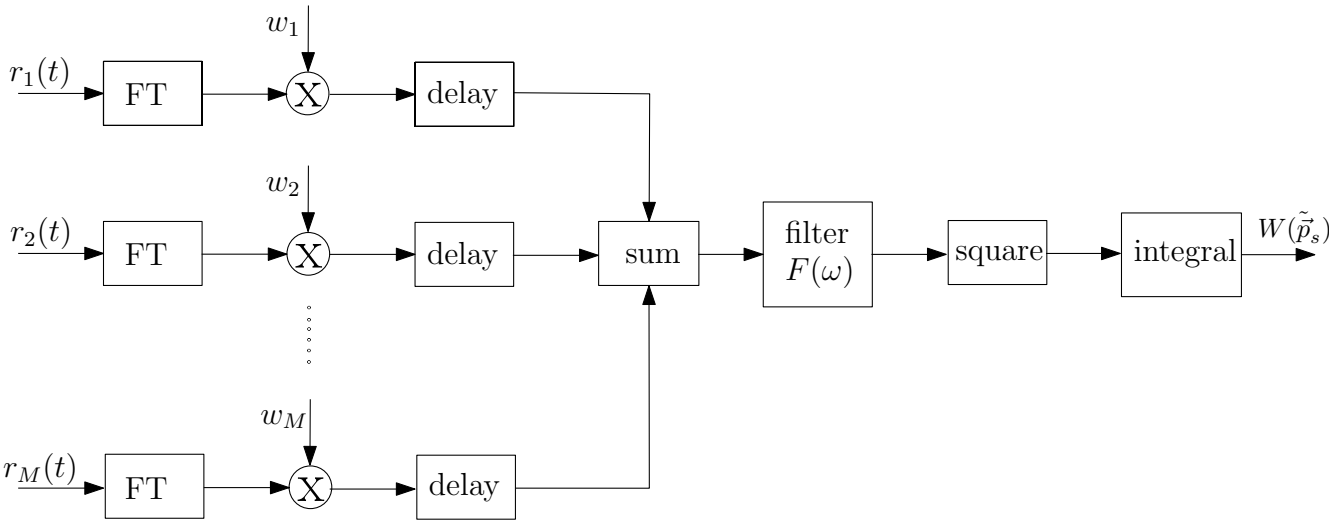

Figure 7.2: One common filter for all input sensor signals

where

$$A(F) = C^A \int_{-\infty}^{\infty} |F(\omega)|^4 \omega^2 S_{ss}(\omega) S_{nn}(\omega) d\omega, \tag{7.8}$$

$$B(F) = C^B \int_{-\infty}^{\infty} |F(\omega)|^4 \omega^2 S_{nn}^2(\omega) d\omega, \tag{7.9}$$

$$D(F) = \sqrt{C^D} \int_{-\infty}^{\infty} |F(\omega)|^2 \omega^2 S_{ss}(\omega) d\omega, \tag{7.10}$$

with $S_{ss}(\omega)$ being the source signal energy spectrum, $S_{nn}(\omega)$ being the noise power spectrum, and

$$C^A = \frac{2}{\pi} \sum_{m=1}^{M} \sum_{l=1}^{M} \sum_{n=1}^{M} \frac{w_m w_l^2 w_n}{d_m d_n} (t_m - t_l)(t_n - t_l), \tag{7.11}$$

$$C^B = 2 \sum_{m=1}^{M} \sum_{l=1}^{M} w_m^2 w_l^2 (t_m - t_l)^2, \tag{7.12}$$

$$C^D = \left[\frac{1}{2\pi} \sum_{m=1}^{M} \sum_{l=1}^{M} \frac{w_m^2 w_l^2}{d_m d_l} (t_m - t_l)^2 \right]^2, \tag{7.13}$$

$$t_m = -\frac{\vec{e} \cdot \vec{p}_m}{d_m c}. \tag{7.14}$$

where $\vec{e}$ denotes the estimation error direction, ϵ the estimation error along $\vec{e}$, and $\sigma_\epsilon^2(F)$ the estimation error variance.

It is seen that $B(F)$ represents the pure noise and $A(F)$ represents a distorted signal, which can also be viewed as a mixed noise of the original noise and signal. These two factors together contribute to the estimation error, and therefore a truly effective filter must address the two components at the same time. Though the EEV has been selected as the measure for performance evaluation, the SNR approach to improve the system performance is also investigated due to its popularity. The results are presented at the end of

the next section.

7.5 Derivation of Suboptimal EEV filters

According to (7.6), it is reasonable to define $F(\omega)$ as a low pass filter with a frequency range of $[-\Omega, \Omega]$, which is the same as the source signal bandwidth so that the noise energy can be limited in order to increase the SNR. Then based on (7.7), several suboptimal filter expressions can be derived after the integration limits in $A(F)$, $B(F)$ and $D(F)$ are changed from $(-\infty, \infty)$ to $[-\Omega, \Omega]$. The term "suboptimal EEV filters", is used here to indicate that these filters do not meet the performance, in terms of minimizing the EEV expression (7.7), of the filter presented in Section 7.6.

Since the value of $\sigma_\epsilon^2(F)$ is determined by the sum of $\frac{A(F)}{D^2(F)}$ and $\frac{B(F)}{D^2(F)}$, the first term to consider is the cross term, $\frac{A(F)}{D^2(F)}$, representing the mixed noise of pure additive Gaussian noise and the signal. The term $\frac{B(F)}{D^2(F)}$, whose numerator represents the pure noise, is left unattended for the moment. Observing $A(F)$, rewrite the denominator $D^2(F)$ as

$$D^2(F) = C^D \left(\int_{-\Omega}^{\Omega} |F(\omega)|^2 \omega \sqrt{S_{ss}(\omega) S_{nn}(\omega)} \omega \sqrt{\frac{S_{ss}(\omega)}{S_{nn}(\omega)}} d\omega \right)^2,$$

then apply Cauchy-Schwarz inequality,

$$D^2(F) \leq C^D \int_{-\Omega}^{\Omega} |F(\omega)|^4 \omega^2 S_{ss}(\omega) S_{nn}(\omega) d\omega \int_{-\Omega}^{\Omega} \omega^2 \frac{S_{ss}(\omega)}{S_{nn}(\omega)} d\omega,$$

where the equality above holds iff

$$|F(\omega)|^2 = \frac{1}{S_{nn}(\omega)}. \tag{7.15}$$

Note that (7.15) is equivalent to the expression of the so-called noise whitening filter given in [13]. Substitute (7.15) into $A(F)/D^2(F)$ and $B(F)/D^2(F)$ to give

$$\frac{A(F)}{D^2(F)} \geq \frac{C^A}{C^D \int_{-\Omega}^{\Omega} \omega^2 \frac{S_{ss}(\omega)}{S_{nn}(\omega)} d\omega}, \tag{7.16}$$

$$\frac{B(F)}{D^2(F)} = \frac{C^B \int_{-\Omega}^{\Omega} |F(\omega)|^4 \omega^2 S_{nn}^2(\omega) d\omega}{C^D \left(\int_{-\Omega}^{\Omega} \omega^2 \frac{S_{ss}(\omega)}{S_{nn}(\omega)} d\omega \right)^2}. \tag{7.17}$$

Though the filter given by (7.15) minimizes term $A(F)/D^2(F)$, the same effect on $B(F)/D^2(F)$ does not occur, since $F(\omega)$ still exists in it. Obviously the noise whitening filter is a suboptimal solution. Substituting (7.15) into $B(F)/D^2(F)$, the corresponding theoretical EEV value can be calculated. Furthermore, when the noise is simply white Gaussian, this filter has no impact on $\sigma_\epsilon^2(F)$ at all.

Similarly, $B(F)/D^2(F)$ can be minimized while $A(F)/D^2(F)$ is left unattended, which gives

$$|F(\omega)|^2 = \frac{S_{ss}(\omega)}{S_{nn}^2(\omega)}, \tag{7.18}$$

the same as the expression for the Eckart filter. Due to its incapability to minimize the term $\frac{A(F)}{D^2(F)}$, it is also a suboptimal solution.

The two suboptimal filters presented so far are both derived through the use of the EEV expression. For applications where a closed-form EEV expression is unavailable, the signal-to-noise ratio provides a good alternative to evaluate and improve the system performance. Maximizing the SNR has been proven an equivalent approach to minimizing directly the EEV in [6], when amplitude weighting is used to optimize the system performance. Hence, before the filter that can minimize the value of (7.7), is derived in the next section, it is desirable to study what the SNR approach can contribute to frequency weighting.

Maximizing the SNR expression given by the right-hand side of (E.4) in Appendix E gives the following filter transfer function:

$$|F_{bound}(\omega)| = \frac{\sum_{m=1}^{M} \frac{w_m}{d_m}}{\sum_{m=1}^{M} w_m^2} \left| \frac{X_s(\omega)}{S_{nn}(\omega)} \right|. \tag{7.19}$$

$F_{bound}(\omega)$ brings an upper bound value for the SNR defined by (E.1). It is equivalent to the Eckart filter when applied to the FWEGS algorithm, since the coefficient $\frac{\sum_{m=1}^{M} \frac{w_m}{d_m}}{\sum_{m=1}^{M} w_m^2}$ in (7.19) is constant for all grid points and cancels out in (7.7).

7.6 Derivation of the Filter Minimizing the EEV

The previous section discussed several filters that can reduce the value of (7.7). In this section, a filter that minimizes (7.7) is derived.

With $F(\omega)$ being defined as a filter transfer function over $\omega \in [-\Omega, \Omega]$, the classical method to obtain the optimal $F(\omega)$ expression is through the minimization of $\sigma_\epsilon^2(F)$, using the calculus of variations technique. It is reasonable to assume that (7.7) is a quasiconvex function that has an unambiguous unique global minimum under certain amount of noise, because the work carried out has been centered around small error analysis, which implies that the total energy function has only one global maximum around the actual peak. Its first and second order conditions are expected to be met when the minimizer is applied.

Let

$$K(\omega) = C^A \omega^2 S_{ss}(\omega) S_{nn}(\omega) + C^B \omega^2 S_{nn}^2(\omega), \tag{7.20}$$

$$L(\omega) = \sqrt{C^D} \omega^2 S_{ss}(\omega), \tag{7.21}$$

$$F_2(\omega) = |F(\omega)|^2. \tag{7.22}$$

thus (7.7) can be written as

$$\sigma_\epsilon^2(F) = \frac{\int_{-\Omega}^{\Omega} K(\omega) F_2^2(\omega) d\omega}{\left[\int_{-\Omega}^{\Omega} L(\omega) F_2(\omega) d\omega\right]^2}. \tag{7.23}$$

Now the problem is formulated to find the variation of the functional given by (7.23) with respect to $F_2(\omega)$. Based on the Euler Equation, a minimizer for $\sigma_\epsilon^2(F)$ can be found, which is defined as the filter that is "optimal" in the context of this work.

The variation of $\sigma_\epsilon^2(F)$ is given by

$$\begin{aligned}\delta\sigma_\epsilon^2 &= 2\int_{-\Omega}^{\Omega}\delta F_2(\omega)\left[K(\omega)F_2(\omega)d\omega\int_{-\Omega}^{\Omega}L(\omega)F_2(\omega)d\omega\right.\\ &\quad\left.-L(\omega)\int_{-\Omega}^{\Omega}K(\omega)F_2^2(\omega)\right]d\omega\Bigg/\left[\int_{-\Omega}^{\Omega}L(\omega)F_2(\omega)d\omega\right]^3. \end{aligned}\tag{7.24}$$

Since $\delta F_2(\omega)$ is arbitrary and continuous, and the filter will be designed as a band pass filter that satisfies $F_2(-\Omega) = F_2(\Omega) = 0$, there has to be

$$K(\omega)F_2(\omega)\int_{-\Omega}^{\Omega}L(\omega)F_2(\omega)d\omega - L(\omega)\int_{-\Omega}^{\Omega}K(\omega)F_2^2(\omega)d\omega = 0.\tag{7.25}$$

Solving (7.25) above and then substituting with (7.20), (7.21) and (7.22) leads to

$$|F(\omega)|^2 = \lambda\frac{S_{ss}(\omega)}{C^A S_{ss}(\omega)S_{nn}(\omega) + C^B S_{nn}^2(\omega)}, \lambda > 0.\tag{7.26}$$

Substitute solution (7.26) into (7.7), λ cancels out. Therefore, λ can be assigned with the value of 1 to simplify calculation when (7.26) is applied to the FWEGS algorithm, where λ is a constant for all grid points.

7.7 Simulation and Performance Comparison

In this section, the performance of each filter is evaluated using Monte Carlo simulation consisting of 1000 iterations using a series of system SNR values. The system SNR is defined as the total energy ratio between the signal and noise across all the sensors, given by (E.1). The EEV values for each filter under each SNR are then plotted as functions of SNR to have each filter performance visualized.

The suboptimal filters derived in Section 7.5 are equivalent to two well-known filters given in the literature: the noise whitening filter and Eckart Filter. The filter developed for optimal time delay estimation in [13] and the simplest form of a traditional matched filter are also chosen to be evaluated and compared here. They have been briefly introduced earlier in Section 7.1. All the filters considered for simulation in this

section are listed below in their spectrum representation:

$$|F_{Mch}| = |X_s(\omega)|, \tag{7.27}$$

$$|F_{Eck}|^2 = S_{ss}(\omega)/S_{nn}^2(\omega), \tag{7.28}$$

$$|F_{Wh}|^2 = 1/S_{nn}(\omega), \tag{7.29}$$

$$|F_{SuboptD}|^2 = \frac{S_{ss}(\omega)/S_{nn}^2(\omega)}{1 + M[S_{ss}(\omega)/(S_{nn}(\omega))]}, \tag{7.30}$$

$$|F_{Opt}|^2 = \frac{S_{ss}(\omega)}{S_{nn}(\omega)\left[C^A S_{ss}(\omega) + C^B S_{nn}(\omega)\right]}, \tag{7.31}$$

where F_{Mch} represents the matched filter, F_{Eck} the Eckart filter, F_{Wh} the noise whitening filter, $F_{SuboptD}$ the filter proposed for optimal time delay estimation in [13] that is modified to be applicable here, and F_{Opt} that minimizes the analytic EEV expression given by (7.7). $F_{SubOptD}$ is chosen to be part of this simulation because it also considers the signal-noise cross term as F_{Opt} does. The difference is that the former adopts the number of sensors while the latter uses theoretically derived coefficients. It is desirable to see if more accurate coefficients can actually make a difference. Only the magnitude of the filters has been considered, since phase information is unknown from the statistically obtained source signal energy spectrum. Also the algorithm given by (7.6) itself just considers energy, which means only the magnitude of $F(\omega)$ is important.

The filter expressions themselves indicate that under white Gaussian noise, the noise whitening filter will bring no performance improvement and the Eckart filter will become equivalent to the matched filter. Hence in this simulation colored noise is used. Its power spectrum is generated using Matlab filter function block and shown in Figure 7.3.

The FWEGS algorithm is motivated for microseismic event localization for the potash mining industry. Hence the source signal used in the simulation, depicted in Figure 7.4, is the energy spectrum obtained statistically [21] using the data collected at Lanigan Potash Mine, Saskatchewan, Canada. The sensor coordinates are indicated as numerically labeled in the mine map given in Figure 7.5, demonstrating that the sensors are not deployed in a typical regular array form of a line or circle as seen in most studies in the literature.

In this simulation, it is assumed that coarse grid searches are already done and a small area that is believed to contain the true source position has been identified, within which finer grid searches are performed. The sampling frequency f_s adopted for this simulation is 1000 Hz; the grid step adopted is 2 m; the error direction $e = [1,\ 0]$; and for each SNR value a Monte Carlo simulation consisting of 1000 iterations is performed.

The simulation results are plotted in Figure 7.6 using four subplots so that the performance differences between the filters can be seen more clearly. The Whitening filter, F_{Wh}, and the filter aiming for optimal delay, $F_{SubOptD}$, yield very close performance as shown in subplots (a) and (b) of Figure 7.6. Neither filter truly brings improvement in this scenario, probably because as they suppress the noise, the signal is also suppressed to almost the same degree, given the noise and signal shapes adopted in this simulation. Filters

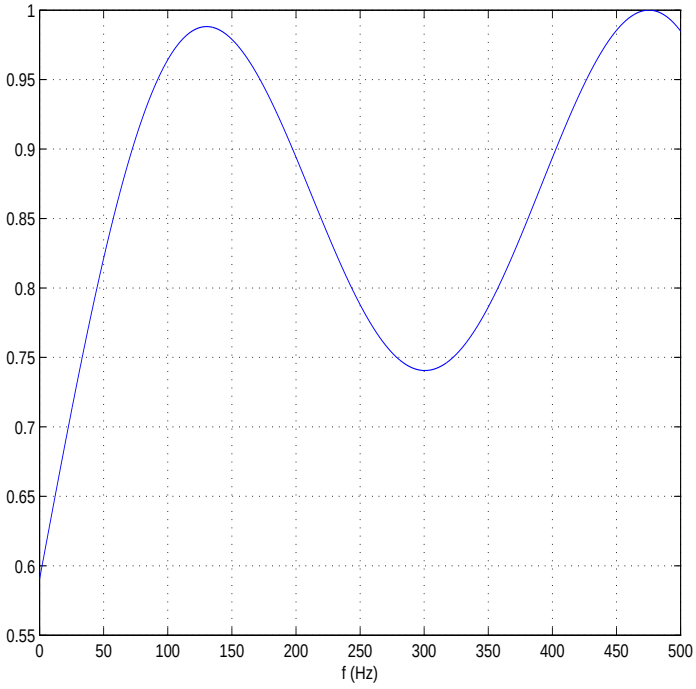

Figure 7.3: Generated Power Spectrum of Random Noise

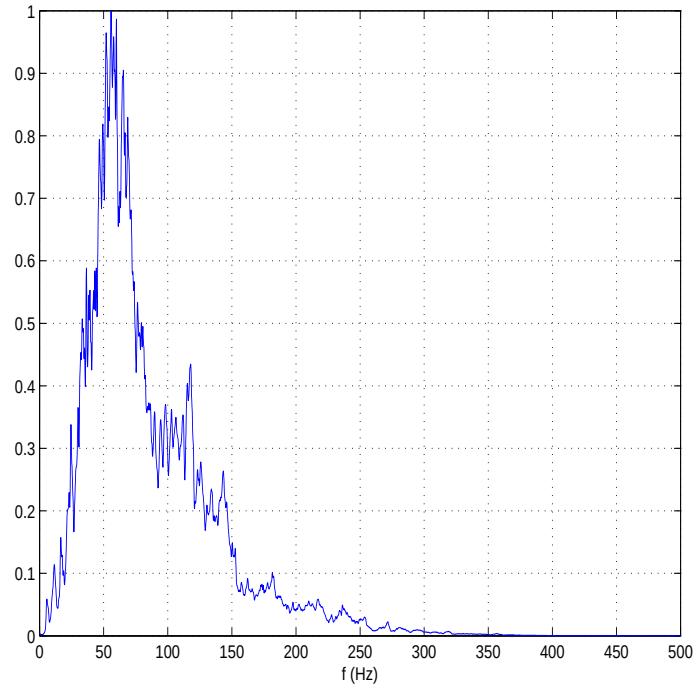

Figure 7.4: Real Statistical Microseismic Event Signal Energy Spectrum

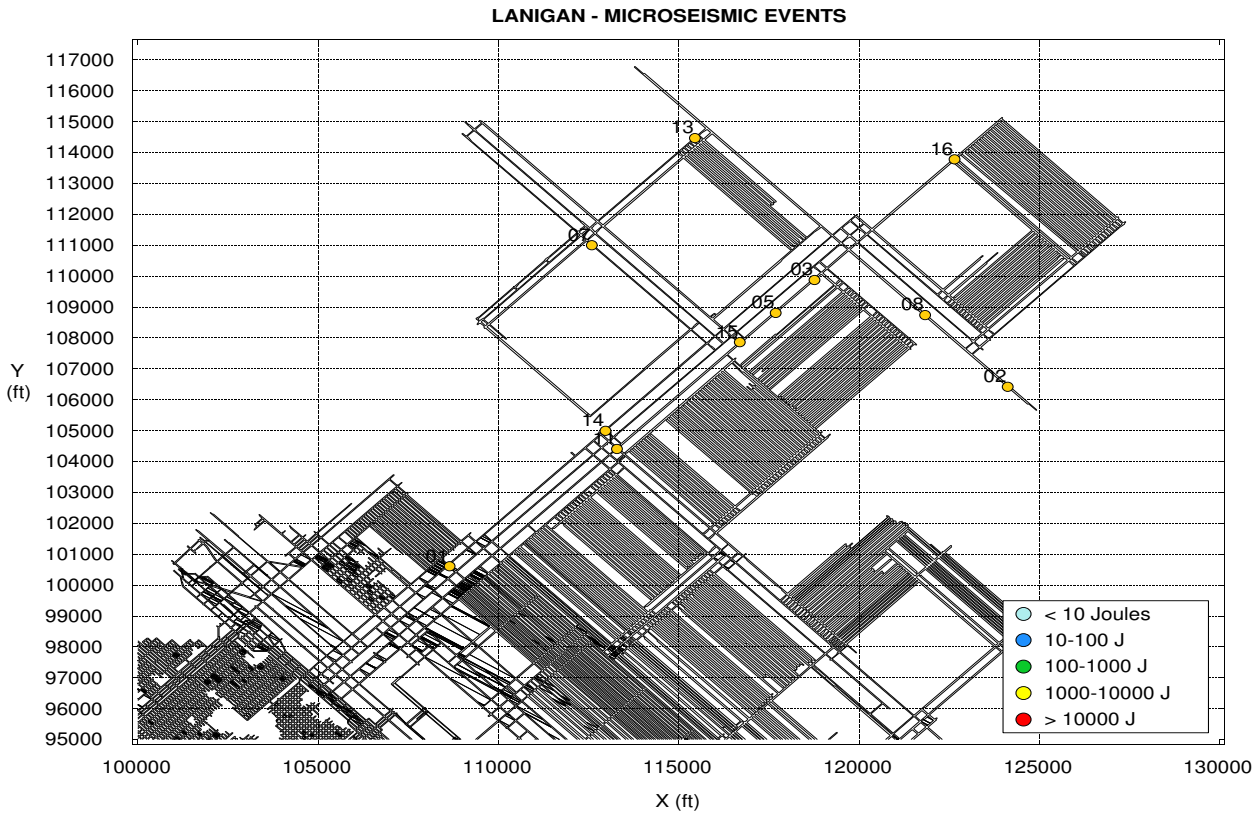

Figure 7.5: Real sensor configuration in a potash mine

F_{Eck}, F_{Mch}, and F_{Opt} demonstrate much better performance as a group compared to F_{Wh} and $F_{SubOptD}$ in subplots (c) and (d) of Figure 7.6. As expected, the filter minimizing the analytic EEV expression given by (7.7) also brings the lowest actual EEV among all the filters considered in this paper, though all of the others converge to similar performance with increasing SNR. This indirectly proves that the analytic EEV derived closely approximates the actual EEV.

To further demonstrate the performance difference between the filter minimizing (7.7) and the other filters considered, the same simulation results are replotted in Figure 7.7 to indicate the percentages by which the EEV values of the other filters are higher than that of F_{Opt}.

It is seen that when the noise is high but still within the range for the filters to work, the performance improvement brought by F_{Opt} can be as high as almost 400%. It is interesting to note that when the SNR gets higher, even though all the filters start performing better, F_{Opt} experiences much greater performance improvement. It is understandable, since neither F_{Eck}, nor F_{Mch} considers as F_{Opt} does, the signal-noise cross term, which can be much more significant relative to the pure noise term.

7.8 Conclusion

Following the work in [6], where the sum of cross-correlation technique and grid search are used for microseismic source localization in the mining industry, this paper has investigated optimal filtering for the application of interest to achieve better location estimate accuracy.

It is demonstrated that when information about the source signal and noise characteristics is available, the use of properly designed filters can help improve the source location estimation accuracy. All the filters derived in previous sections are low pass filters with different magnitudes to limit the noise energy and/or magnify the source signal energy so that a better system SNR can be achieved.

It is intuitive to assume that when the system SNR is maximized, the performance will be optimized as well. This has been theoretically proven to be true in the case of amplitude weighting in [6], which shows that maximizing the defined SNR is equivalent to minimizing the $\sigma_\epsilon(F)$. However, the SNR and EEV approaches are not theoretically equivalent for frequency weighting investigated in this paper. There can be more than one definition for the system SNR, but only one for the EEV. Hence maximizing an SNR expression does not necessarily produce a minimum for the source location EEV.

When applied to the FWEGS algorithm described by (7.6), $F_{bound}(\omega)$ derived by dealing with the SNR is equivalent to F_{Eck}. This filter expression does not address the signal-noise cross term that also contributes to the estimation error variance. In cases that the value of the signal-noise cross term is not negligible but comparable to or even much greater than that of the pure noise term, F_{Opt} is expected to outperform all the other filters considered, which has been demonstrated in the simulation results. For scenarios where the signal-noise cross term is small enough to be ignored, F_{Opt} may yield similar performance as F_{Eck}. Stating that minimizing the EEV is equivalent to maximizing the SNR in case of frequency weighting may seem true in effect in certain scenarios, but not generally true.

According to the theoretical derivation and simulation results obtained in this paper, suboptimal filters such as F_{Wh} and F_{Mch}, may sometimes serve as good replacements for F_{Opt} because they are comparatively simpler to implement. F_{Wh} may be adopted when the noise statistical characteristics are known but the source signal is unknown, while F_{Mch} may be adopted when the noise is white and the source signal energy spectrum is known.

Acknowledgments

The authors would like to thank TRlabs, Saskatoon and the Potash Corporation of Saskatchewan Inc., for supporting this research.

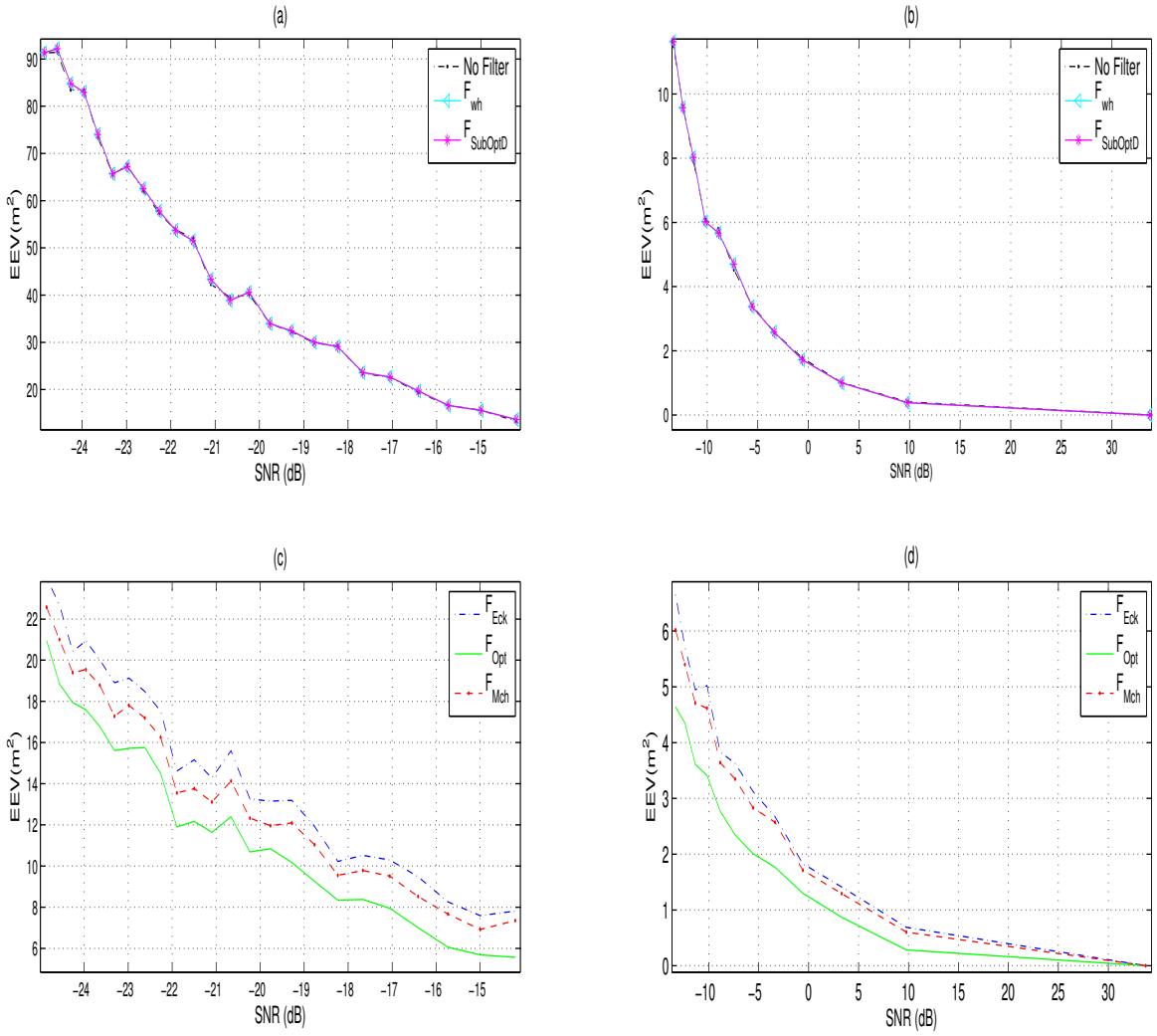

Figure 7.6: FWEGS Performance With Various Filter Applied

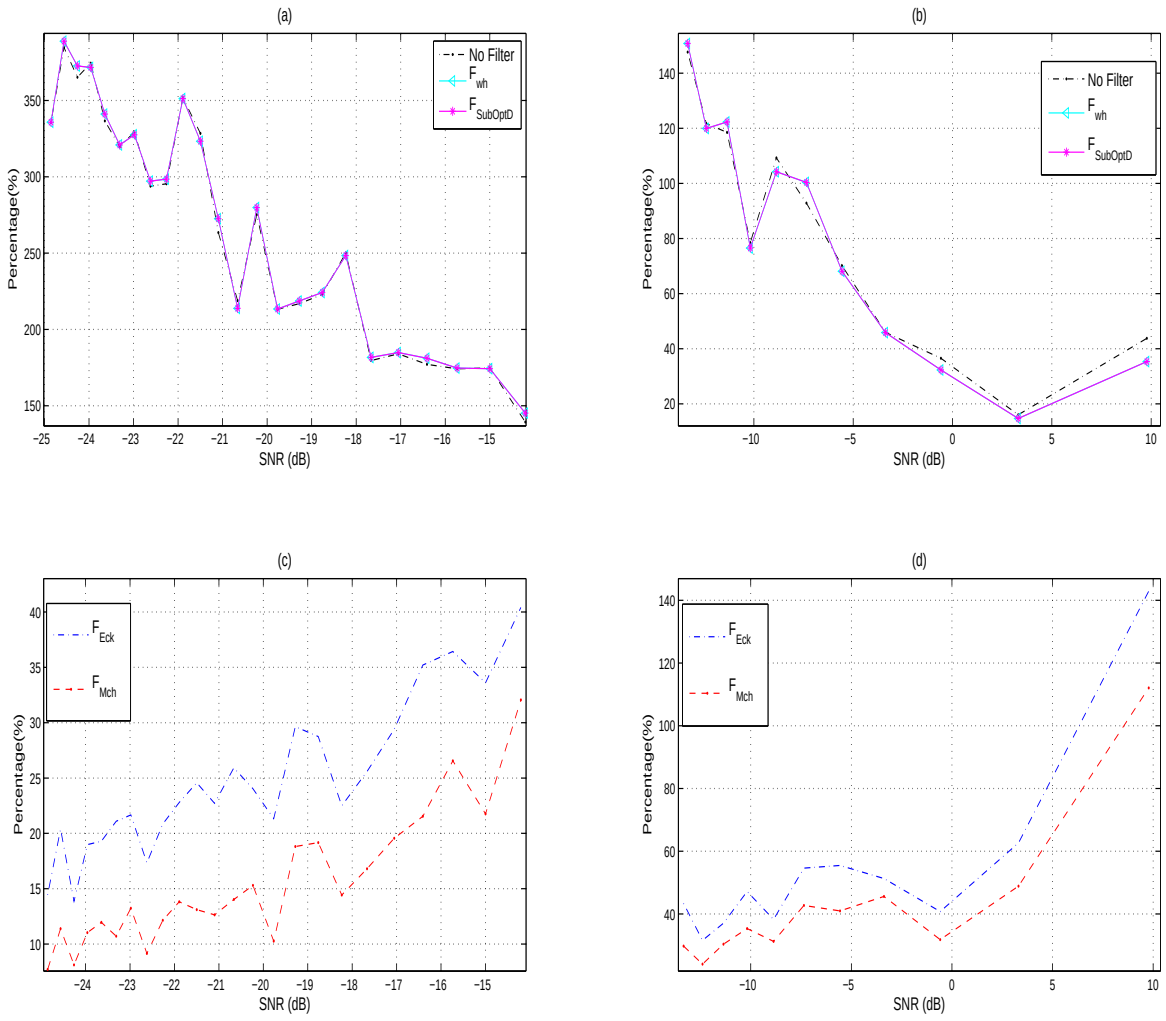

Figure 7.7: Optimal Filter Performance Improvement over the other Filters

D Derivation of $\sigma_\epsilon^2(F)$

Based on the work of [5], similar analysis can be performed in the frequency domain. With arguments of the hypothesized source position $\tilde{\vec{p}}_s$ and a filter $F(\omega)$ whose time domain expression denoted $f(t)$, a total energy function in the frequency domain can be defined as

$$\begin{aligned}\mathbf{W}(\tilde{\vec{p}}_s) &= \frac{1}{2\pi}\int_{-\infty}^{\infty}\left|F(\omega)\sum_{m=1}^{M} w_m\left(e^{j\omega(\tilde{\tau}_m-\tau_m)}\frac{X_s(\omega)}{d_m}+e^{j\omega\tilde{\tau}_m}X_{n_m}(\omega)\right)\right|^2 d\omega \\ &= \mathbf{W}_s+\mathbf{W}_{sn}+\mathbf{W}_n, \end{aligned} \tag{D.1}$$

where

$$\mathbf{W}_s = \frac{1}{2\pi}\int_{-\infty}^{\infty}|F(\omega)|^2\sum_{m=1}^{M}\sum_{l=1}^{M}\frac{\tilde{w}_m\tilde{w}_l}{d_m d_l}e^{j\omega(\Delta\tau_m-\Delta\tau_l)}|X_s(\omega)|^2 d\omega; \tag{D.2}$$

$$\begin{aligned}\mathbf{W}_{sn} &= \frac{1}{2\pi}\int_{-\infty}^{\infty}|F(\omega)|^2\sum_{m=1}^{M}\sum_{l=1}^{M}\left[\frac{w_m w_l}{d_m}e^{j\omega(\Delta\tau_m-\tilde{\tau}_l)}X_s(\omega)X_{n_l}^*(\omega)\right. \\ &\quad \left.+\frac{w_m w_l}{d_l}e^{j\omega(\tilde{\tau}_m-\Delta\tau_l)}X_s^*(\omega)X_{n_m}(\omega)\right]d\omega; \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}|F(\omega)|^2\sum_{m=1}^{M}\sum_{l=1}^{M}\frac{w_m w_l}{d_m}\left[e^{j\omega(\Delta\tau_m-\tilde{\tau}_l)}X_s(\omega)X_{n_l}^*(\omega)\right. \\ &\quad \left.+e^{j\omega(\tilde{\tau}_l-\Delta\tau_m)}X_s^*(\omega)X_{n_l}(\omega)\right]d\omega; \end{aligned} \tag{D.3}$$

$$\mathbf{W}_n = \frac{1}{2\pi}\int_{-\infty}^{\infty}|F(\omega)|^2\sum_{m=1}^{M}\sum_{l=1}^{M}\tilde{w}_m\tilde{w}_l e^{j\omega(\tilde{\tau}_m-\tilde{\tau}_l)}X_{n_m}(\omega)X_{n_l}^*(\omega)d\omega. \tag{D.4}$$

with $\Delta\tau_m=\tilde{\tau}_m-\tau_m$.

Similar to the analyzing procedure carried out in [5], it is found that the EEV of the FAWEGS algorithm $\sigma_\epsilon^2(F)$ can be expressed in the frequency domain as

$$\sigma_\epsilon^2(F)=\frac{A(F)+B(F)}{D^2(F)}, \tag{D.5}$$

where $A(F)=E\left[\left(\frac{d\mathbf{W}_{sn}}{d\tilde{\epsilon}}\right)^2\right]\Big|_{\tilde{\vec{p}}_s=\vec{p}_s}$, $B(F)=E\left[\left(\frac{d\mathbf{W}_n}{d\tilde{\epsilon}}\right)^2\right]\Big|_{\tilde{\vec{p}}_s=\vec{p}_s}$, and $D^2(F)=\frac{d^2\mathbf{W}_s}{d\tilde{\epsilon}^2}\Big|_{\tilde{\vec{p}}_s=\vec{p}_s}$.

Finding that

$$\frac{d\Delta\tau_m}{d\tilde{\epsilon}} = \frac{d\tilde{\tau}_m}{d\tilde{\epsilon}} = -\frac{\vec{e}\cdot\vec{p}_m}{cd_m}, \tag{D.6}$$

$$\text{and} \qquad \Delta\tau_m|_{\tilde{\vec{p}}_s=\vec{p}_s} = 0, \tag{D.7}$$

it is easy to obtain that

$$\begin{aligned} D(F) &= \left.\frac{d^2\mathbf{W}_s}{d\tilde{\epsilon}^2}\right|_{\tilde{\vec{p}}_s=\vec{p}_s} \\ &= \frac{1}{2\pi}\int_{-\infty}^{\infty}|F(\omega)|^2\omega^2 S_{ss}(\omega)d\omega\sum_{m=1}^{M}\sum_{l=1}^{M}\frac{\tilde{w}_m w_l}{d_m d_l}(t_m-t_l)^2. \end{aligned} \tag{D.8}$$

where $S_{ss}(\omega) = |X_s(\omega)|^2$ is the energy spectrum of the source signal and

$$t_m = -\frac{\vec{e}\cdot\vec{p}_m}{d_m c}. \tag{D.9}$$

Next the term $A(F)$ is to be found. Given (D.3) and (D.9), there is

$$\begin{aligned} \left(\frac{d\mathbf{W}_{sn}}{d\tilde{\epsilon}}\right)^2\Bigg|_{\tilde{\vec{p}}_s=\vec{p}_s} &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(\omega_1)|^2|F(\omega_2)|^2(j\omega_1 j\omega_2) \\ &\quad \sum_{m=1}^{M}\sum_{l=1}^{M}\sum_{n=1}^{M}\sum_{k=1}^{M}\frac{w_m w_l w_n w_k}{d_m d_n}(t_m-t_l)(t_n-t_k)\times \\ &\quad \Big(X_s(\omega_1)X_s(\omega_2)X^*_{n_l}(\omega_1)X^*_{n_n}(\omega_2)e^{-j\omega_1\tau_l-j\omega_2\tau_n} \\ &\quad -X_s(\omega_1)X^*_s(\omega_2)X^*_{n_l}(\omega_1)X_{n_n}(\omega_2)e^{j\omega_2\tau_n-j\omega_1\tau_l} \\ &\quad -X^*_s(\omega_1)X_s(\omega_2)X_{n_l}(\omega_1)X^*_{n_n}(\omega_2)e^{j\omega_1\tau_l-j\omega_2\tau_n} \\ &\quad +X^*_s(\omega_1)X^*_s(\omega_2)X_{n_l}(\omega_1)X_{n_n}(\omega_2)e^{j\omega_1\tau_l+j\omega_2\tau_n}\Big)d\omega_1 d\omega_2. \end{aligned} \tag{D.10}$$

Apparently that the expected values of the above four terms inside the double integrals are only nonzero when $n = l$. Hence taking the expected value of $\left(\frac{d\mathbf{W}_{sn}}{d\tilde{\epsilon}}\right)^2$ brings

$$\begin{aligned} A(F) &= E\left[\left(\frac{d\mathbf{W}_{sn}}{d\tilde{\epsilon}}\right)^2\right]\Bigg|_{\tilde{\vec{p}}_s=\vec{p}_s} \\ &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(\omega_1)|^2|F(\omega_2)|^2(j\omega_1 j\omega_2)\times \\ &\quad \sum_{m=1}^{M}\sum_{l=1}^{M}\sum_{n=1}^{M}\frac{w_m w_l^2 w_n}{d_m d_n}(t_m-t_l)(t_n-t_l)\times \\ &\quad \Big(X_s(\omega_1)X_s(\omega_2)E[X^*_{n_l}(\omega_1)X^*_{n_l}(\omega_2)]e^{-j(\omega_1+\omega_2)\tau_l} \\ &\quad -X_s(\omega_1)X^*_s(\omega_2)E[X^*_{n_l}(\omega_1)X_{n_l}(\omega_2)]e^{-j(\omega_1-\omega_2)\tau_l} \\ &\quad -X^*_s(\omega_1)X_s(\omega_2)E[X_{n_l}(\omega_1)X^*_{n_l}(\omega_2)]e^{j(\omega_1-\omega_2)\tau_l} \\ &\quad +X^*_s(\omega_1)X^*_s(\omega_2)E[X_{n_l}(\omega_1)X_{n_l}(\omega_2)]e^{j(\omega_1+\omega_2)\tau_l}\Big)d\omega_1 d\omega_2. \end{aligned} \tag{D.11}$$

Substitute $E[X_{n_l}(\omega_1)X_{n_l}(\omega_2)]$, $E[X_{n_l}(\omega_1)X^*_{n_l}(\omega_2)]$, $E[X^*_{n_l}(\omega_1)X_{n_l}(\omega_2)]$, and $E[X^*_{n_l}(\omega_1)X^*_{n_l}(\omega_2)]$ utilizing (12-68), (12-71) and (12-75) in [22], and noticing that $X^*_s(\omega) = X_s(-\omega)$, there is

$$\begin{aligned}
A(F) &= E\left[\left(\frac{d\mathbf{W}_{sn}}{d\tilde{\epsilon}}\right)^2\right]\Bigg|_{\tilde{\vec{p}}_s=\vec{p}_s} \\
&= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(\omega_1)|^2|F(\omega_2)|^2(-\omega_1\omega_2)\times \\
&\quad \sum_{m=1}^{M}\sum_{l=1}^{M}\sum_{n=1}^{M}\frac{w_m w_l^2 w_n}{d_m d_n}(t_m - t_l)(t_n - t_l)\times \\
&\quad \Big[2\pi S_{nn}(\omega_1)\delta(-\omega_1-\omega_2)X_s(\omega_1)X_s(\omega_2)e^{j(-\omega_1-\omega_2)\tau_l} \\
&- \quad 2\pi S_{nn}(\omega_1)\delta(-\omega_1+\omega_2)X_s(\omega_1)X^*_s(\omega_2)e^{j(-\omega_1+\omega_2)\tau_l} \\
&- \quad 2\pi S_{nn}(\omega_1)\delta(\omega_1-\omega_2)X^*_s(\omega_1)X_s(\omega_2)e^{j(\omega_1-\omega_2)\tau_l} \\
&+ \quad 2\pi S_{nn}(\omega_1)\delta(\omega_1+\omega_2)X^*_s(\omega_1)X^*_s(\omega_2)e^{j(\omega_1+\omega_2)\tau_l}\Big]d\omega_1 d\omega_2 \\
&= \frac{2}{\pi}\int_{-\infty}^{\infty}|F(\omega)|^4\omega^2 S_{ss}(\omega)S_{nn}(\omega)\times \\
&\quad \sum_{m=1}^{M}\sum_{l=1}^{M}\sum_{n=1}^{M}\frac{w_m w_l^2 w_n}{d_m d_n}(t_m - t_l)(t_n - t_l)d\omega. \qquad \text{(D.12)}
\end{aligned}$$

Finally, $B(F)$ is to be determined. Given (D.9), take the derivative of (D.4) with respective to $\tilde{\epsilon}$, square and then take the expected value to give

$$\begin{aligned}
\left(\frac{d\mathbf{W}_n}{d\tilde{\epsilon}}\right)^2\Bigg|_{\tilde{\vec{p}}_s=\vec{p}_s} &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(\omega_1)|^2|F(\omega_2)|^2 j\omega_1 j\omega_2\times \\
&\quad \sum_{m=1}^{M}\sum_{l=1}^{M}\sum_{n=1}^{M}\sum_{k=1}^{M}w_m w_l w_n w_k\times \\
&\quad (t_m - t_l)(t_n - t_k)e^{j\omega_1(\tau_m-\tau_l)}e^{j\omega_2(\tau_n-\tau_k)}\times \\
&\quad E\left[X_{n_m}(\omega_1)X^*_{n_l}(\omega_1)X_{n_n}(\omega_2)X^*_{n_k}(\omega_2)\right]d\omega_1 d\omega_2. \qquad \text{(D.13)}
\end{aligned}$$

The four random variables in the expectation of (D.13) are all zero mean and jointly normally distributed

[22] [23], and we get

$$\begin{aligned} B(F) &= E\left[\left(\frac{d\mathbf{W}_n}{d\tilde{\epsilon}}\right)^2\right]\Big|_{\tilde{\vec{p}}_s=\vec{p}_s} \\ &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(\omega_1)|^2|F(\omega_2)|^2(-\omega_1\omega_2)\times \\ &\quad \sum_{m=1}^{M}\sum_{l=1}^{M}\sum_{n=1}^{M}\sum_{k=1}^{M}\frac{w_m w_l w_n w_k}{}(t_m-t_l)(t_n-t_k)\times \\ &\quad \Big(E[X_{n_m}(\omega_1)X^*_{n_l}(\omega_1)]E[X_{n_n}(\omega_2)X^*_{n_k}(\omega_2)]+ \\ &\quad E[X_{n_m}(\omega_1)X_{n_n}(\omega_2)]E[X^*_{n_l}(\omega_1)X^*_{n_k}(\omega_2)]+ \\ &\quad E[X_{n_m}(\omega_1)X^*_{n_k}(\omega_2)]E[X_{n_n}(\omega_2)X^*_{n_l}(\omega_1)]\Big)\times \\ &\quad e^{j\omega_1(\tau_m-\tau_l)}e^{j\omega_2(\tau_n-\tau_k)}d\omega_1 d\omega_2. \end{aligned} \tag{D.14}$$

Since whenever $E[X_{n_m}(\omega_1)X^*_{n_l}(\omega_1)]E[X_{n_n}(\omega_2)X^*_{n_k}(\omega_2)]$ is nonzero, $(t_m-t_l)(t_n-t_k)$ is zero, the first term of (D.14) always equals zero. The second term is only nonzero when $m=n$ and $l=k$ and the third term is only nonzero when $m=k$ and $l=n$.

Utilizing (12-68), (12-71) and (12-75) in [22] again, there is

$$\begin{aligned} B(F) &= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|F(\omega_1)|^2|F(\omega_2)|^2(-\omega_1\omega_2)\times \\ &\quad \Big[2\pi S_{nn}(\omega_1)\delta(\omega_1+\omega_2)2\pi S_{nn}(\omega_1)\delta(-\omega_1-\omega_2)\times \\ &\quad \sum_{m=n=1}^{M}\sum_{l=k=1}^{M}w_m^2 w_l^2(t_m-t_l)^2 e^{j(\omega_1+\omega_2)(\tau_m-\tau_l)} \\ &+ \quad 2\pi S_{nn}(\omega_1)\delta(\omega_1-\omega_2)2\pi S_{nn}(\omega_1)\delta(-\omega_1+\omega_2)\times \\ &\quad \sum_{m=k=1}^{M}\sum_{l=n=1}^{M}w_m^2 w_l^2(t_m-t_l)(t_l-t_m)e^{j(\omega_1-\omega_2)(\tau_m-\tau_l)}\Big]d\omega_1 d\omega_2 \\ &= 2\int_{-\infty}^{\infty}|F(\omega)|^4\omega^2 S_{nn}^2(\omega)d\omega\sum_{m=1}^{M}\sum_{l=1}^{M}w_m^2 w_l^2(t_m-t_l)^2. \end{aligned} \tag{D.15}$$

Now substitute (7.11), (7.12), (7.13) into terms $A(F)$, $B(F)$ and $D(F)$ given by (D.12), (D.15) and (D.8) respectively, the EEV expression given by (D.5) is arrived

$$\sigma_\epsilon^2(F) = \frac{C^A\int_{-\infty}^{\infty}|F(\omega)|^4\omega^2 S_{ss}(\omega)S_{nn}(\omega)d\omega + C^B\int_{-\infty}^{\infty}|F(\omega)|^4\omega^2 S_{nn}^2(\omega)d\omega}{C^D\left[\int_{-\infty}^{\infty}|F(\omega)|^2\omega^2 S_{ss}(\omega)d\omega\right]^2}. \tag{D.16}$$

E Filter Obtained by Maximizing SNR

With $s_0(t)$ and $n_m(t)$ representing the source and noise signal, a signal-to-noise energy ratio, SNR_{eng}, can be defined as

$$SNR_{eng} = \frac{\int_{-\infty}^{\infty} |f(t) * \sum_{m=1}^{M} \frac{w_m}{d_m} s(t-\tau_m)|^2 dt}{E\{\int_{-\infty}^{\infty} |f(t) * \sum_{m=1}^{M} w_m n_m(t)|^2 dt\}}, \tag{E.1}$$

where w_m represents the true amplitude weight and $*$ denotes convolution. This SNR definition results naturally from the algorithm given by (7.6) that considers the total energy of the received sensor signals. However, it is very difficult to obtain an analytic expression for the filter transfer function by maximizing (E.1) directly. Given the non-negativity of the integrated function in the numerator of (E.1), maximizing (E.1) is mathematically equivalent to maximizing the ratio of the numerator value at any specific time t_0 to the same denominator value. Hence the following can replace (E.1) to be maximized,

$$SNR_p = \frac{\left| f(t) * \sum_{m=1}^{M} \frac{w_m}{d_m} s(t-\tau_m) \right|^2_{t=t_0}}{E\{\int_{-\infty}^{\infty} |f(t) * \sum_{m=1}^{M} w_m n_m(t)|^2 dt\}}, \tag{E.2}$$

Rewrite (E.2) using frequency domain representation, apply properties of Fourier Transform to the numerator and Parseval's theorem to the denominator, there is

$$\begin{aligned} & SNR_p \\ = & \frac{\frac{1}{4\pi^2} \left| \int_{-\infty}^{\infty} F(\omega) \sum_{m=1}^{M} \frac{w_m}{d_m} X_s(\omega) e^{j\omega(-\tau_m)} e^{j\omega t} d\omega \right|^2_{t=t_0}}{\frac{1}{2\pi} \int_{-\infty}^{\infty} |F(\omega)|^2 \sum_{m=1}^{M} \sum_{l=1}^{M} E\{X_{n_m}(\omega) X^*_{n_l}(\omega)\} w_m w_l d\omega} \\ = & \frac{\left| \int_{-\infty}^{\infty} F(\omega) \sum_{m=1}^{M} \frac{w_m}{d_m} X_s(\omega) e^{j\omega(-\tau_m)} e^{j\omega t} d\omega \right|^2_{t=t_0}}{2\pi \int_{-\infty}^{\infty} |F(\omega)|^2 S_{nn}(\omega) \sum_{m=1}^{M} w_m^2 d\omega}, \end{aligned} \tag{E.3}$$

where $F(\omega)$, $X_s(\omega)$ and $X_{n_m}(\omega)$ are the Fourier transforms of the filter transform function, source and noise signal respectively.

(E.3) resembles the more traditional system SNR definition, which can be further modified into

$$SNR_p < SNR_{p1} = \frac{\left| \int_{-\infty}^{\infty} F(\omega) \sum_{m=1}^{M} \frac{w_m}{d_m} |X_s(\omega)| d\omega \right|^2}{2\pi \int_{-\infty}^{\infty} |F(\omega)|^2 S_{nn}(\omega) \sum_{m=1}^{M} w_m^2 d\omega}. \tag{E.4}$$

Rearrange the numerator of (E.4) and then apply Cauchy-Schwarz Inequality to it,

$$SNR_{p1} \leq \frac{\int_{-\infty}^{\infty} \left| F(\omega) \sqrt{S_{nn}(\omega) \sum_{m=1}^{M} w_m^2} \right|^2 d\omega \times \int_{-\infty}^{\infty} \left| \frac{|X_s(\omega)| \sum_{m=1}^{M} \frac{w_m}{d_m}}{\sqrt{S_{nn}(\omega) \sum_{m=1}^{M} w_m^2}} \right|^2 d\omega}{2\pi \int_{-\infty}^{\infty} |F(\omega)|^2 S_{nn}(\omega) \sum_{m=1}^{M} w_m^2 d\omega}. \tag{E.5}$$

The equality in (E.5) holds iff

$$\left| F(\omega) \sqrt{S_{nn}(\omega) \sum_{m=1}^{M} w_m^2} \right|^2 = \left| \frac{|X_s(\omega)| \sum_{m=1}^{M} \frac{w_m}{d_m}}{\sqrt{S_{nn}(\omega) \sum_{m=1}^{M} w_m^2}} \right|^2, \tag{E.6}$$

which leads to

$$|F(\omega)| = \left| \frac{X_s(\omega)}{S_{nn}(\omega)} \right| \frac{\sum_{m=1}^{M} \frac{w_m}{d_m}}{\sum_{m=1}^{M} w_m^2}. \tag{E.7}$$

The filter transfer function given by (E.7) maximizes the right-hand side of (E.4), representing an upper bound of the system SNR defined by (E.1).

References

[1] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 24, pp. 320–327, August 1976.

[2] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. on Acoustics, Speech and Signal Processing*, vol. 35, pp. 1223–1225, August 1987.

[3] J. O. Smith and J. S. Abel, "The spherical interpolation method of source localization," *IEEE J. of Oceani. Eng.*, vol. 12, pp. 246–252, January 1987.

[4] J. Nelder and R. Mead, "A simplex method for function minimization," *Computer Journal*, vol. 7, pp. 308–313, 1965.

[5] B. L. Daku and E. J. Salt, "Error analysis of a localization algorithm for finite-duration events," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1024–1034, March 2007.

[6] S. Li and B. L. Daku, "Optimal amplitude weighting for near-field passive source localization," *IEEE Trans. on Signal Processing*, vol. 59, pp. 6175 – 6185, August 2011.

[7] M. G. Christensen, J. H. Jensen, andreas Jakobsson, and S. H. Jensen, "On optimal filter designs for fundamental frequency estimation," *IEEE Signal Processing Letters*, vol. 15, pp. 745–748, 2008.

[8] M. Basin, J. Perez, and R. Martinez-Zuniga, "Alternative optimal filter for linear state delay systems," *International Journal of Adaptive Control and Signal Processing*, vol. 20, pp. 509–517, April 2006.

[9] T. Tung, K. Yao, D. Chen, R. Hudson, and C. Reed, "Source localization and spatial filtering using wideband MUSIC and maximum power beamforming for multimedia applications," in *1999 IEEE Workshop on Signal Processing Systems: Design and Implementation*, (Taipei, Taiwan), pp. 625–634, SIPS 99, 1999.

[10] D. Korompis, K. Yao, and F. Lorenzelli, "Broadband maximum energy array with user imposed spatial and frequency constraints," in *Proceedings of ICASSP-94*, vol. iv, pp. IV529–IV532, ICASSP-94, April 1994.

[11] S. Alliney and C. Morandi, "Digital image registration using projections," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. PAPI-8, pp. 222–233, March 1986.

[12] W.K.Pratt, "Correleation techniques for image registration," *IEEE Trans. on Aerosp. Electron. Syst.*, vol. AES-10, pp. 353–358, May 1974.

[13] W. R. Hahn and S. A. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Trans. on Information Theory*, vol. IT-19, pp. 508–614, September 1973.

[14] G. Bensen, M.H.Ritzwoller, M.P.Barmin, A.L.Levshin, F.Lin, M.P.Moschetti, N.M.Shapiro, and Y.Yang, "Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements," *Geophys. J. Int.*, vol. 169, pp. 1239–1260, Jan 2007.

[15] M. Schimmel, E. Stutzmann, F. Ardhuin, and J. Gallart, "Using instantaneous phase coherence for signal extraction from ambient noise data at a local to a global scale," *Geophys. J. Int.*, vol. 184, pp. 494–506, 2011.

[16] C. D. Mcgillem and M. Svedlow, "Image registration error variance as a measure of overlay quality," *IEEE Trans. on Geoscience Electronics*, vol. GE-14, pp. 44–49, January 1976.

[17] C. D. Mcgillem and M. Svedlow, "Optimum filter for minimization of image registration error variance," *IEEE Trans. on Geoscience Electronics*, vol. GE-15, pp. 257–259, October 1977.

[18] T. L. Steding and F. W. Smith, "Optimum filter for image registration," *IEEE Trans. on Aerospace and Electronic Systems*, vol. AES-15, pp. 849–859, November 1979.

[19] P. M. J. van der Zwet, M. Nettesheim, J. J. Gerbrands, and J. H. C. Reiber, "Derivation of optimal filters for the detection of coronary arteries," *IEEE Trans. on Medical Imaging*, vol. 17, pp. 108–120, February 1998.

[20] C. Eckart, "Optimal rectifier systems for the detection of steady signals," tech. rep., Univ. California, Scripps Inst. Oceanography, Marine Physical Lab. Rep SI0 12692, 1952. Univ. California, Scripps Inst. Oceanography, Marine Physical Lab., SI0 12692, SI0 Ref 52-11.

[21] A. Errington, B. L. Daku, D. Dodds, and A. Prugger, "Characterization of source signal in the potash mines," in *Proceedings of CCECE05*, (Saskatoon, SK, Canada), May 2005.

[22] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, 3rd ed., 1991.

[23] L. Isserlis, "On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables," *Biometrika*, vol. 12, pp. 134–139, 1918.

8 Summary and Conclusions

The problem of passive acoustic source localization was investigated in this thesis. It was motivated by the specific application of microseismic event localization in the vicinity of potash mines in the province of Saskatchewan, Canada. The application is characterized by the involvement of a source signal with a small time-bandwidth product and sensors that are arranged in the near field. The major topics of this thesis are the selection, evaluation and performance optimization of a source localization algorithm with better consistency but without introducing too much extra calculation load for the application of interest. In addition, the research results are applicable to a wide range of applications where the sum of cross-correlation technique is used for signal processing.

The first task accomplished was presented in Chapter 4, which is the selection of a one-step estimation algorithm for the application of interest. When the source signal lasts long enough for the TDOAs to be estimated with an acceptable accuracy, classical two-step source localization algorithms are good choices. After the TDOAs are estimated, the source location can be further estimated using analytic techniques, such as the hyperbolic intersection, or numeric techniques, such as the simplex search method. However, the source signal under consideration has a small time-bandwidth product and requires some different treatment. Therefore, three representative algorithms given in the literature for acoustic source localization were evaluated and compared in Chapter 4. The results led to the choice of the EGS algorithm. Compared with two-step location estimators, the EGS algorithm avoids the likelihood of introducing extra error into the final source location estimate by eliminating the step of TDOA estimation. Compared with the AML estimator, which is also a one-step estimator but processes the original data in the frequency domain, the EGS estimator suffers no performance degradation when the data samples are limited, or when the signal is non-stationary. The EGS algorithm can also be implemented more easily in either software or hardware when compared with the AML algorithm, because the EGS algorithm does not require a signal processing module to perform the Fourier transform.

The second achievement is the development of the corresponding Cramér-Rao lower bound expressions using the Fisher information matrix and was presented in Chapter 5. These CRB expressions provide a critical tool to evaluate the estimators investigated in this study. The derivation starts with the assumption of the signal and noise being time-continuous. This is to avoid the complex matrix calculation required to

diagonalize the covariance matrix of a CGN random process when time-discrete data are used. Assuming the time-continuous signal form also makes it consistent with the work presented in [1]. Many works in the literature focus only on white Gaussian noise in order to ensure that the Fourier transform of the noise is orthogonal along the frequency axis. In this thesis, Karhunen-loève decomposition is adopted, so that a continuous random process given in the time domain can be mapped into a series of independent discrete random variables in the frequency domain. Therefore, non-white Gaussian noise can be easily accommodated, where the mathematical handling is similar to that of white Gaussian noise. The derivation procedure is also applicable to non-stationary Gaussian noise problems. However, explicit CRB expressions have only been reached for stationary Gaussian noise cases in this thesis.

The EGS algorithm was evaluated in Chapter 6, using the CRB expressions derived in Chapter 5, and was shown to have room for improvement. Note that the signal attenuation at different sensors is not considered by the original EGS algorithm. Hence the first optimization strategy investigated was the utilization of the amplitude difference across the sensors, so that a more accurate source location could be estimated. The algorithm proposed in Chapter 6 was therefore, named the Amplitude-Weighted EGS algorithm. The noise model adopted in Chapter 6 was simplified from the general model presented in Chapter 3 and assumes identically and independently distributed white Gaussian noise at each sensor. This assumption enables the algorithm to weigh each sensor signal with a real positive value calculated according to the amplitude of the received signal. The absolute value of the received signal strength at any specific sensor is not critical, but its relative relationship to everyone else is. The AWEGS EEV was obtained after including weights in the EGS EEV derivation procedure presented in [1]. A closed-form expression for the set of weights, optimal in terms of minimizing the analytic AWEGS estimation error, was derived. In reality, an analytic EEV expression of an estimator is not always available. Therefore, performance optimization through the use of the SNR was also investigated. A system SNR was defined for the application and a closed-form expression for the set of weights optimal in terms of maximizing this SNR was then reached. Note that the two approaches, minimizing the AWEGS EEV and maximizing the SNR, are equivalent in the realm of amplitude weighting under the assumption of i.i.d white Gaussian noise at each sensor

The optimal weight expression derived suggests that the optimal weights themselves have to be estimated first. A common practice is to assume all the weights to be "1" to initialize the search. Then an estimated source location is solved. With this estimated source location, a set of weight values closer to be "optimal" can be calculated and reapplied to the next iteration. This procedure will repeat until certain predefined threshold is hit. Grid search itself is already a heavy burden to the EGS algorithm in terms of calculation. Having to do a number of iterations to obtain the optimal weight values makes the burden even heavier. These additional calculations required compromise much of the benefit brought by the proposed amplitude-weighted algorithm. Therefore, an alternate non-iterative weighting method was proposed and brought optimal performance with just one iteration for a given grid search. Monte Carlo simulation indicates that

the proposed non-iterative amplitude-weighted algorithm is desirable as expected.

It is worth noting that the AWEGS algorithm proposed is an intuitively and straightforwardly constructed unbiased estimator for passive source localization, over complete statistics of observed random data. It utilizes all the available source location information carried in both the time delays and amplitude differences between sensor signals. This means it is the *efficient* estimator for near-field scenarios. It is also known that if an Maximum-likelihood(ML) estimator exists, it is the *efficient* estimator. Hence the AWEGS algorithm is essentially the ML estimator for the application of interest, though not constructed using the model-based approach required by the ML method.

Amplitude weighting can be applied to any applications utilizing the energy in the sum of sensor signals, regardless of the type of the source signal. It is a good candidate when the source signal and noise are unknown and the noise can be modeled as WGN. However, in certain situations, the source signal and noise are both known and a dedicated signal processing technique for them is desired.

To address this, a frequency weighting strategy was proposed in Chapter 7 to yield more accurate source location estimate. First, the frequency-weighted EGS EEV was derived according to the selected way of applying a filter to the original EGS algorithm. Then several sub-optimal filters, matching several well-known classical filters given in the literature, were derived based on the FWEGS EEV expression obtained. These filters are called "sub-optimal" because they don't meet the performance of the filter that minimizes the analytic FWEGS EEV expression developed in this work. The latter one is named the "optimal filter" in the context of this work. All the filters derived in this chapter are low pass filters with different envelopes to limit the noise energy and/or magnify the source signal energy so that a better system SNR can be achieved. It is intuitive to assume that when the system SNR is maximized, the performance will be optimized as well. This has been theoretically proven to be true in the case of amplitude weighting in Chapter 6, which shows that maximizing the system SNR given by (A.3) is equivalent to minimizing the EEV given by (6.5). However, the SNR and EEV approaches are not theoretically equivalent for frequency weighting, and this has been shown in Chapter 7, due to the fact that there can be more than one definition for the SNR but only one for the EEV. Therefore stating that minimizing the EEV is equivalent to maximizing the SNR in case of frequency weighting may seem true in certain scenarios, but not generally true. F_{opt} outperforms all the other filters considered when the value of signal-noise cross term is not negligible but comparable to or even greater than that of the pure noise term, as expected. For scenarios where the signal-noise cross term is small enough to be ignored, F_{opt} may yield similar performance as F_{eck}.

According to the theoretical derivation and simulation results obtained in Chapter 7, suboptimal filters such as F_{wh} and F_{mch}, can sometimes be good replacements for F_{opt} because they are comparatively simpler to implement in practice. F_{wh} can be adopted when the noise statistical characteristics are known but the source signal is unknown, while F_{mch} can be adopted when the noise is white and the source signal energy

spectrum is known.

To summarize the above concisely: 1. One-step source localization is more robust compared with two-step localization; 2. The EGS algorithm performed over the original time domain signals without any extra frequency domain processing can achieve good estimation results with limited samples; 3. The analytic source location EEV is minimized in the realm of amplitude weighting when a set of optimal weight values derived is applied to the AWEGS algorithm, which is an efficient estimator; 4. The amplitude weighting technique can be realized in a non-iterative way, which greatly reduces the extra calculation burden that is required by the traditional iterative way of implementation; 5. The analytic source location EEV is minimized in the realm of frequency weighting when the filter F_{Opt} developed is applied to the FWEGS algorithm. 6. Various shapes of source and signal spectrum produce various relations between the pure noise and the signal-noise cross term, both of which contribute to the EEV value. Among all the filters investigated in this study, F_{Opt} is the only one that addresses the two terms correctly, which results in more consistent localization performance.

Possible future work can be performed to investigate: 1. The effect of noise and signal spectrum envelopes on the source location estimation error; 2. Sensor arrangement optimization for near-field source localization; 3. Actual implementation of the FWEGS algorithm with various filters of interest applied.

References

[1] B. L. Daku and E. J. Salt, "Error analysis of a localization algorithm for finite-duration events," *IEEE Trans. on Signal Processing*, vol. 55, pp. 1024–1034, March 2007.