NETWORK AND SENSOR MANAGEMENT

FOR MULTIPLE SENSOR EMITTER

LOCATION SYSTEM

BY

XI HU

BEng, Liaoning University, China, 1998
MEng, Beijing Jiaotong University, China, 2001
MS, Binghamton University, Binghamton, New York, 2007

DISSERTATION

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate School of
Binghamton University
State University of New York
2008

Accepted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate School of
Binghamton University
State University of New York
2008

October 8, 2008

Mark L. Fowler, Department of Electrical and Computer Engineering, Binghamton University

Eva Wu, Department of Electrical and Computer Engineering, Binghamton University

Xiaohua Li, Department Electrical and Computer Engineering, Binghamton University

ABSTRACT

In order for estimating the location of a passive emitter, signal data must be collected at a multitude of sensors and the sensors must cooperate to achieve the task. Our goal is to achieve network-wide optimization over a large number of simultaneously deployed sensors to enable more efficient and effective cooperation within the network of sensors.

This dissertation covers the following aspects: i) The emitter location estimation accuracy is related with many items, giving an overall review of how the measurements quality (such as the accuracies of TDOA and FDOA) and sensors' navigation data (such as position and velocity) will affect the estimation errors, and developing the relationship among those aspects; (ii) Since the accuracy of parameter estimation is related with the signal model, exploiting the importance of deciding the signal model used in location estimation problem and giving out the results of different models; (iii) To save the system energy and reduce computation latency, developing various methods to select and pair a subset of sensors to satisfy the system requirements; (iv) Based on the relationship between estimation accuracy and sensors' navigation data, discussing the probability of computing the next optimal state; (v) Since sensors' navigation data along with uncertainty, exploiting the sensors' navigation data errors effects on least square estimation and giving solution to mitigate these errors. The results of this dissertation will provide a systematic means for addressing network management and sensor management issues across the spectrum of sensor network.

ACKNOWLEDGEMENTS

I wish to acknowledge the guidance and advising of my advisor, Professor Mark L. Fowler. I appreciate everything that he has done in helping my dissertation along, and for offering support and encouraging during my studying. I would also like to thank Professor Eva Wu for her helpful discussion and guidance concerning my research. Additionally, I thank two committee members Professor Xiaohua Li and Professor Harold W. Lewis for their service on my dissertation committee and their reviewing my dissertation.

I would like to thank my parents for their never ending love. Especially, I appreciate my husband for all he has done to me; he is always supportive of me. Without any of you, this dissertation would never have been possible.

Also thanks to my friends at Binghamton University, they make my life here wonderful.

CHAPTERS

LIST OF FIGURES

1 Introduction

Estimating the location of a passive emitter has been a research issue for decades ([6]-[21]). The estimation procedure has two steps. First: estimate sensors' received signal parameters, such as arriving time, frequency, angle, phase, or the energy. Second: use the parameters estimated in the first stage to estimate the emitter location. These sensors are located at some vehicles or aircrafts, and the unmanned vehicles are becoming more and more popular. In order for wireless sensor networks to exploit a signal, signal data must be collected at a multitude of sensors and the sensors must cooperate to achieve the task.

Our general interest is in achieving network-wide optimization over a large number of simultaneously deployed sensors to enable more efficient and effective cooperation within the network of sensors. Fisher information can be used to assess the data quality across multiple sensors to manage the network of sensors to optimize the location accuracy subject to communication constraints. For emitter location it is well-known that the geometry between sensors and the target plays a key role in determining the location accuracy. Furthermore, the deployed sensors have different data quality. Given these two factors, it is no trivial matter to do the network and sensor management as mentioned above.

Whatever methods used to estimate the emitter location will come with the estimation errors. The estimation accuracy is related with many aspects: the sensors' received signal data quality, the sensors' positions and velocities, and the related geometrical properties among sensors and emitter. For examples, [16] gives the "one-sigma-width" to characterize the emitter location accuracy, which is based on only one signal parameter measurement. But for TDOA/FDOA (time difference of arrival/frequency difference of

arrival) localization problem, we need at least two (for two dimensional system) signal parameters measurements to locate the emitter. [8] links the conventional accuracy measures to the moments and products of inertia of a mass configuration and gives some special geometry examples. We exploit the relationship between the emitter location accuracy and all other accuracies of relative measurements and geometrical aspects in general cases. It gives an overview of how all these aspects will affect the final accuracy, and the trade off among them.

For the given sensors network, it is desirable to use the complete set of sensor resources to do the task. However, that will result in an excessive data computation and communication in the network. And the main constraints of wireless sensor network are the limited on-board energy and limited channel resources. Many approaches have been proposed in the past to satisfy the network resource requirements, such as routing, sleep modes, low-power electronics, etc. For examples: one data compression method was proposed by Professor Fowler and Dr. Mo Chen in Dr. Mo Chen's dissertation [23]. Sensor selection is one of the solutions to save system energy which is to select a subset of sensors to achieve the requirement. Using the information theory to do the sensor selection was proposed in [36][37]. A new method is proposed in this dissertation: Fisher information based Sensor Selection and Pairing, which is more efficient and accurate.

Our work is to manage the given set of sensors to satisfy the emitter location accuracy demanding. We propose various approaches to this problem and discuss trade-offs among them. For examples: One method assumes that the sensors have pre-paired and share their data between these pairs; sensor selection then consists of selecting pairs to optimize performance while meeting constraints on number of pairs selected; Another method

consists of optimally determining pairings as well as selections of pairs with or without sensor sharing.

To estimate the emitter location, we need some signal's measured/estimated parameters from sensors' received data, and we also need the sensors' navigation data, such sensors' positions and velocities. In practice, the positions and velocities of sensors can not be known exactly. A closed form solution to take the receiver error into account was given in [40], which uses two steps estimation methods to solve the problems. We exploit an expression about how these navigation data will affect the estimation accuracy and give out an error mitigation processing, which will play an important role in the decrease of estimation deviation. Our method is more simple and with less assumptions compare with others' methods.

Also if we can maneuver the sensors, based on the current information we have about the sensors and the data quality, we can decide the next optimal states of sensors (sensors' positions and velocities) within the next research sets. It is called the maneuver of sensors or trajectory planning. We proposed the basic idea about the next optimal states. And developing solutions for some sub-optimal problems.

The contributions of this dissertation include: i) Giving an overall review of how the measurements quality (such as the accuracies of TDOA and FDOA) and sensors' navigation data (such as position and velocity) will affect the estimation errors, and developing the relationship among those aspects; (ii) Exploiting the importance of deciding the signal model used in location estimation problem, and giving out the different results for different models on Fisher information calculation, maximum

likelihood estimator and etc.; (iii) Developing various methods to select and pair a subset of sensors to satisfy the system requirements, such as saving system energy and reduce computation latency; (iv) Exploiting the sensors' navigation data errors effects on calculation Fisher information and least square estimation, and giving solution to mitigate these errors. (v) Discussing the probability of computing the next optimal state, and giving some results on sub-optimal problems; The results of this dissertation will provide the engineer with a systematic means for addressing network management and sensor management issues across the spectrum of sensor network.

2 Parameter Estimation

Emitter localization is to estimate the emitter location by some parameters of the signal sent by the emitter. There are two stages parameter estimations: one is the signal parameter estimation, the other one is the emitter location estimation. In this chapter we introduce the basic concepts, properties and two estimation methods of parameter estimation we used in this dissertation.

2.1 Signal Model

Parameter estimation is to estimate a parameter through some measured data, which depend on the unknown parameter. Assume we measured N samples data as $\mathbf{x} = \{x[0], \quad x[1], \quad \cdots, \quad x[N-1]\}$, the parameter estimator can be written as

$$\hat{\theta}(\mathbf{x}) = g(x[0], \quad x[1], \quad \cdots, \quad x[N-1]) \tag{2-1}$$

where g is the function we used to estimate the parameter θ.

To estimate the parameter of a signal, we need a signal model first, which specifies the relationship between the unknown parameter and received data. The data received are random because of the noise coming with it, so we can describe it by its probability density function (PDF). There are two different scenarios related with the deterministic property of the parameter. (1) If the parameter to be estimated is deterministic, then the PDF of received signal is parameterized by the unknown parameter as $p(\mathbf{x};\theta)$. The estimation is called classical estimation, which we chose in this dissertation. (2) The other scenario is that the unknown parameter is a random variable and we have a prior knowledge about it, which is $p(\theta)$. The parameter to be estimated is viewed as one

realization of the random variable θ. The joint PDF of measured data and unknown parameter can be written as $p(\mathbf{x},\theta) = p(\mathbf{x}\,|\,\theta)p(\theta)$, where $p(\mathbf{x}\,|\,\theta)$ is a conditional PDF of $\mathbf{x}$ conditioned on θ. The estimation is called Bayesian method.

For both deterministic and random variable cases, the parameter estimation is an optimal procedure to determine the unknown parameter via the PDF of the measured data. The PDF may also depend on other parameters assumed known. One of this dissertation's tasks is to exploit the importance of the assumed known parameters and optimal them.

Assume that a signal with an unknown parameter θ is observed in noise as

$$x[n] = s[n;\theta] + w[n] \quad n = 0,1,\ldots,N-1 \tag{2-2}$$

The dependence of $s[n;\theta]$ on θ is assumed known. The vector form of (2-2) can be written as

$$\mathbf{x} = \mathbf{s}(\theta) + \mathbf{w} \tag{2-3}$$

For the received radio frequency (RF) signal we are dealing with, the received noise is normally assumed Gaussian noise with zero mean and covariance matrix $\mathbf{C}_{\mathbf{w}}$. We will give the PDF of $\mathbf{x}$ at different scenarios.

1. Deterministic Parameter and Deterministic Signal Model

If both θ and is $s[n]$ are deterministic, the PDF of $\mathbf{x}$ is functionally depends on θ as

$$p(\mathbf{x};\theta) = \frac{1}{(2\pi)^{N/2}\left|\mathbf{C}_{\mathbf{w}}\right|^{1/2}} \exp\left\{-\frac{1}{2}\left[\mathbf{x}-\mathbf{s}(\theta)\right]^{T}\mathbf{C}_{\mathbf{w}}^{-1}\left[\mathbf{x}-\mathbf{s}(\theta)\right]\right\} \tag{2-4}$$

2. Deterministic Parameter and Random Signal Process

If θ is deterministic, but $s[n]$ is a random process and assumed having Gaussian distribution with zero mean and the variance of it depends on θ. Then the PDF of $\mathbf{x}$ is functionally depends on θ as

$$p(\mathbf{x};\theta)=\frac{1}{(2\pi)^{N/2}\left|\mathbf{C}_{\mathbf{x}}(\theta)\right|^{1/2}}\exp\left\{-\frac{1}{2}\mathbf{x}^{T}\mathbf{C}_{\mathbf{x}}^{-1}(\theta)\mathbf{x}\right\} \tag{2-5}$$

where $\mathbf{C}(\theta)$ means the covariance matrix of $\mathbf{x}$ is a function of θ, and it depends on both the signal and noise.

Why we need to consider the signal model to be deterministic or random for both deterministic unknown parameter cases? Since the key distinction between these two scenarios will drive difference results in finding an optimal estimator for the unknown parameter. One of this dissertation's contributions is to exploit the importance to make a correct assumption on the signal model. We discuss it in chapter 3.

3. Random Parameter and Deterministic Signal Model

In this case, the unknown parameter is assumed random with prior known PDF $p(\theta)$. The conditional PDF of observed data conditioned on θ is

$$p(\mathbf{x}\mid\theta)=\frac{1}{(2\pi)^{N/2}\left|\mathbf{C}_{\mathbf{w}}\right|^{1/2}}\exp\left\{-\frac{1}{2}\left[\mathbf{x}-\mathbf{s}(\theta)\right]^{T}\mathbf{C}_{\mathbf{w}}^{-1}\left[\mathbf{x}-\mathbf{s}(\theta)\right]\right\} \tag{2-6}$$

Then the joint PDF of $\mathbf{x}$ and θ is

$$p(\mathbf{x},\theta)=\frac{1}{(2\pi)^{N/2}\left|\mathbf{C}_{\mathbf{w}}\right|^{1/2}}\exp\left\{-\frac{1}{2}\left[\mathbf{x}-\mathbf{s}(\theta)\right]^{T}\mathbf{C}_{\mathbf{w}}^{-1}\left[\mathbf{x}-\mathbf{s}(\theta)\right]\right\}\cdot p(\theta) \tag{2-7}$$

4. Random Parameter and Random Signal Process

Both the parameter and signal are random. $s[n]$ assumed having Gaussian distribution with zero mean and the variance of it depends on θ. Then the joint PDF of $\mathbf{x}$ and θ is

$$p(\mathbf{x},\theta)=\frac{1}{(2\pi)^{N/2}\left|\mathbf{C}_{\mathbf{x}}(\theta)\right|^{1/2}}\exp\left\{-\frac{1}{2}\mathbf{x}^{T}\mathbf{C}_{\mathbf{x}}^{-1}(\theta)\mathbf{x}\right\}\cdot p(\theta) \tag{2-8}$$

We will not discuss more about the random parameter cases in this dissertation, since we assume the emitter location is deterministic.

2.2 Cramer-Rao Lower Bound and Fisher Information

An estimator $\hat{\theta}(\mathbf{x})$ is a random variable since it is a function of random variables $\mathbf{x}$. To assess how accurate it is, we need to calculate some characters of a random variable, such as the mean and variance.

- **Bias of an Estimator**

The bias of an estimator is the difference between the true value of the parameter and the expectation value of the estimator as

$$\mathrm{b}(\theta) = \mathrm{E}\left\{\hat{\theta}\right\} - \theta \tag{2-9}$$

Unbiased estimator is the one with zero bias.

- **Mean Square Error (MSE) of an Estimator**

The expectation of the square of estimation error, defined as

$$\begin{aligned} \mathrm{mse}(\hat{\theta}) &= \mathrm{E}\left\{(\hat{\theta} - \theta)^2\right\} \\ &= \mathrm{var}(\hat{\theta}) + \mathrm{b}^2(\hat{\theta}) \end{aligned} \tag{2-10}$$

From (2-10), the MSE is compose of both variance and bias of $\hat{\theta}$. The normally optimal estimator is the one that has the minimum MSE (MMSE).

- **Minimum Variance Unbiased (MVU) Estimator**

A MVU estimator is the one that has the minimum variance among all the unbiased $\hat{\theta}$. For the unbiased estimator, $\mathrm{b}(\theta) = 0$, then $\mathrm{mse}(\hat{\theta}) = \mathrm{var}(\hat{\theta})$, MVU is also MMSE.

The minimum variance of any unbiased estimator of θ is called Cramer-Rao Lower Bound (CRLB) of θ. CRLB provides a useful lower bound on the variance of any

unbiased estimator. If the variance of an estimator equals CRLB for each possible value of θ, then it is the MVU estimator.

- **CRLB, Scalar Parameter [1]**

If it is assumed that the PDF $p(\mathbf{x};\theta)$ satisfies the "regularity" condition

$$\mathrm{E}\left[\frac{\partial \ln p(\mathbf{x};\theta)}{\partial\theta}\right]=0 \quad \textit{for all } \theta \tag{2-11}$$

where the expectation is taken with respect to $p(\mathbf{x};\theta)$. Then the CRLB of θ can be calculated as

$$\mathrm{CRLB}(\theta)=\frac{1}{-\mathrm{E}\left[\frac{\partial^2 \ln p(\mathbf{x};\theta)}{\partial\theta^2}\right]}=\frac{1}{\mathrm{I}(\theta)} \tag{2-12}$$

where the derivative is evaluated at the true value of θ and the expectation is taken with respect to $p(\mathbf{x};\theta)$. $\mathrm{I}(\theta)$ is called the Fisher information (FI) of θ.

The variance of any unbiased estimator $\hat{\theta}(\mathbf{x})$ must satisfy

$$\mathrm{var}(\hat{\theta})\geq \mathrm{CRLB}(\theta) \tag{2-13}$$

So far we talked about scalar parameter case, we now extend the results to the vector parameter case. $\boldsymbol{\theta}=[\theta_1,\cdots,\theta_p]^T$ is the vector to be estimated and $s[n;\boldsymbol{\theta}]$ is the signal which parametered by $\boldsymbol{\theta}$.

- **CRLB, Vector Parameter [1]:**

If the PDF $p(\mathbf{x};\boldsymbol{\theta})$ satisfies the "regularity" condition

$$\mathrm{E}\left[\frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}\right]=\mathbf{0} \quad \textit{for all } \boldsymbol{\theta} \tag{2-14}$$

where the expectation is taken with respect to $p(\mathbf{x};\boldsymbol{\theta})$. Then the Fisher information matrix (FIM) of $\boldsymbol{\theta}$ can be calculated as

$$\left[\mathbf{I}(\boldsymbol{\theta})\right]_{ij} = -\mathrm{E}\left[\frac{\partial^2 \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i \partial\theta_j}\right] \quad 1 \le i, j \le p \tag{2-15}$$

where the derivative is evaluated at the true value of $\boldsymbol{\theta}$ and the expectation is taken with respect to $p(\mathbf{x};\boldsymbol{\theta})$. The CRLB matrix of $\boldsymbol{\theta}$ is

$$\mathbf{CRLB}(\boldsymbol{\theta}) = \mathbf{I}^{-1}(\boldsymbol{\theta}) \tag{2-16}$$

The covariance matrix of any unbiased estimator $\hat{\boldsymbol{\theta}}(\mathbf{x})$ must satisfy

$$\mathbf{C}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \mathrm{var}(\hat{\theta}_1) & \cdots & \mathrm{cov}(\hat{\theta}_1, \hat{\theta}_p) \\ \vdots & \ddots & \vdots \\ \mathrm{cov}(\hat{\theta}_1, \hat{\theta}_p) & \cdots & \mathrm{var}(\hat{\theta}_p) \end{bmatrix} \ge \mathbf{CRLB}(\boldsymbol{\theta}) \tag{2-17}$$

"≥" means $\mathbf{C}(\hat{\boldsymbol{\theta}}) - \mathbf{CRLB}(\boldsymbol{\theta})$ is a semi-positive definite matrix, then

$$\mathrm{var}(\hat{\theta}_i) = [\mathbf{C}(\hat{\boldsymbol{\theta}})]_{ii} \ge [\mathbf{CRLB}(\boldsymbol{\theta})]_{ii} \quad i = 1, 2, ..., p \tag{2-18}$$

Since the Fisher information of the unknown parameter is calculated for the derivative of the PDF of observed data, therefore FI depends on the sensitivity of the PDF on the unknown parameter. The more sensitive the PDF is influenced by the unknown parameter, the larger the FI, the smaller the CRLB and the better we could estimate it, and vice versa. The PDF of $\mathbf{x}$ also depends on the parameters we assumed known. So if the signal model is fixed, we can increase the sensitivity by modify the known parameters. This is the motivation of this dissertation, because Fisher information captures the entire essential trade-offs embedded in the estimation problem.

- **CRLB for the general Gaussian case**

In the case of Gaussian observation assume that $\mathbf{x} \sim N(\boldsymbol{\mu}(\boldsymbol{\theta}), \mathbf{C}(\boldsymbol{\theta}))$, where $\boldsymbol{\mu}(\boldsymbol{\theta})$ is the $N \times 1$ mean vector and $\mathbf{C}(\boldsymbol{\theta})$ is the $N \times N$ covariance matrix, both of them depend on $\boldsymbol{\theta}$. Then the PDF is

$$p(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}\det^{1/2}[\mathbf{C}(\boldsymbol{\theta})]}\exp\left\{-\frac{1}{2}\left[\mathbf{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\right]^T \mathbf{C}^{-1}(\boldsymbol{\theta})\left[\mathbf{x}-\boldsymbol{\mu}(\boldsymbol{\theta})\right]\right\} \quad (2\text{-}19)$$

The FIM is given by [1]

$$\left[\mathbf{I}(\boldsymbol{\theta})\right]_{ij} = \left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right]^T \mathbf{C}^{-1}(\boldsymbol{\theta})\left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta})}{\partial\theta_i}\right] + \frac{1}{2}tr\left[\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_i}\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_j}\right] \quad (2\text{-}20)$$

The FI of each item is composed by two items. The first one depends on mean and covariance and the second one is only related with covariance. We will discuss (2-20) in more details later.

2.3 Maximum Likelihood Estimator

The MVU estimator may not exist or even it exists but maybe not obvious. So sometimes we may use an approximately optimal estimator. The maximum likelihood estimator (MLE) is such an approximately MVU estimator. And we use MLE to estimating TDOA/FDOA in this dissertation, which we will discuss in more details in next chapter. In this section, we introduce what MLE is, how to find MLE and the most important property of it.

$p(\mathbf{x};\theta)d\mathbf{x}$ is the probability of measured $\mathbf{x}$ for a given θ. So for fixed $\mathbf{x}$, the $\hat{\theta}$ that maximize the probability could be the true value of θ. In Figure 1, the Y-axis is the value of $p(\mathbf{x};\theta)$ evaluated at given $\mathbf{x}=\mathbf{x}_0$ and possible value of θ, X-axis is the possible values of θ. Since $p(\mathbf{x}_0;\theta_1)$ has the largest value of $p(\mathbf{x}_0;\theta)$, it is more likely that θ_1 is the true value of θ if $\mathbf{x}=\mathbf{x}_0$ is observed. Or we can say θ_1 maximize the value of $p(\mathbf{x}_0;\theta)$, which is called the likelihood function.

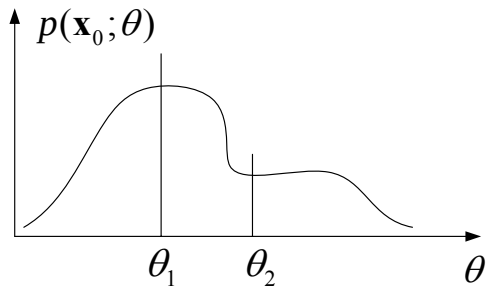

Figure 1 Likelihood function of unknown parameter evaluated on fixed measured data

Definition of MLE: The MLE for a parameter is the value of θ that maximize $p(\mathbf{x};\theta)$ for $\mathbf{x}$ fixed, where $p(\mathbf{x};\theta)$ is called the likelihood function.

The general analytical procedure to find the MLE is:

(1) Find the log-likelihood function: $\ln p(\mathbf{x};\theta)$;

(2) Differentiate $\ln p(\mathbf{x};\theta)$ with respect to θ and set to 0 as $\partial \ln p(\mathbf{x};\theta)/\partial\theta = 0$;

(3) Solve for θ value that satisfies the zero equation.

The definition of scalar parameter MLE is easy to be carried over to the vector parameter case as the vector $\hat{\boldsymbol{\theta}}_{ML}$ that satisfies: $\partial p(\mathbf{x};\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \mathbf{0}$, where

$$\frac{\partial p(\mathbf{x};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} = \begin{bmatrix} \dfrac{\partial p(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_1} \\ \vdots \\ \dfrac{\partial p(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_p} \end{bmatrix} \tag{2-21}$$

The MLE is largely used since for any given $p(\mathbf{x};\boldsymbol{\theta})$ MLE always exists, even if there is no explicit solution we can always find an optimal one by numerical method, and its asymptotically optimal property makes MLE as the optimal estimator.

Asymptotic Property of MLE [1]: If the PDF $p(\mathbf{x};\boldsymbol{\theta})$ satisfies the "regularity" condition as (2-14), then the MLE of $\boldsymbol{\theta}$ is asymptotically Gaussian distributed according to

$$\hat{\boldsymbol{\theta}}_{ML} \overset{a}{\sim} N(\boldsymbol{\theta}, \mathbf{I}^{-1}(\boldsymbol{\theta})) \tag{2-22}$$

where "asymptotically distributed" means for large data measured. The MLE is asymptotically unbiased and asymptotically attains the CRLB, therefore it is asymptotically efficient and optimal, or asymptotically MVU.

In the TDOA/FDOA localization problem, this asymptotical property plays an important role in estimating the emitter location which we will discuss later.

2.4 Least Square Estimator

The MLE needs PDF $p(\mathbf{x};\boldsymbol{\theta})$ known, and normally we assumed it has Gaussian distribution for solving simplicity. But if we do not know the distribution exactly or the PDF is complicated to be simplified, we need the estimators not based on PDF to do the job. Least square estimator (LSE) is one of the estimators that are not statistically based. LSE does not need a PDF model but do need a deterministic signal model. So far we discussed the optimal estimator is the MVU which is unbiased and has minimum variance. LSE uses different criterion to find the optimal solution. It minimizes the difference between the observed data $x[n]$ and the generated noiseless data $s[n;\hat{\boldsymbol{\theta}}]$ as shown in Figure 2, $s[n;\boldsymbol{\theta}]$ is the true noiseless signal parametered by unknown parameter vector $\boldsymbol{\theta}$, $x[n]$ is the perturbed measured data, $s[n;\hat{\boldsymbol{\theta}}]$ is the data generated by estimated parameter $\hat{\boldsymbol{\theta}}$, and $\varepsilon[n]$ is the estimation error also called estimation cost.

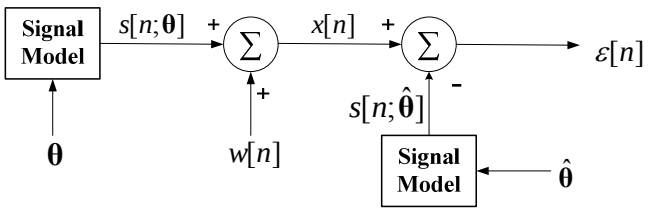

Figure 2 Signal model and least square approach

Definition of LSE: The LSE for a vector parameter is the value of $\boldsymbol{\theta}$ that minimizes the cost function $J(\boldsymbol{\theta})$ which is defined as

$$J(\boldsymbol{\theta}) = \sum_{n=0}^{N-1} \left(x[n] - s[n;\boldsymbol{\theta}] \right)^2 = \left[\mathbf{x} - \mathbf{s}(\boldsymbol{\theta}) \right]^T \left[\mathbf{x} - \mathbf{s}(\boldsymbol{\theta}) \right] \tag{2-23}$$

and

$$\hat{\boldsymbol{\theta}}_{LS} = \arg\min_{\boldsymbol{\theta}} \left\{ J(\boldsymbol{\theta}) \right\} \tag{2-24}$$

- **Weighted Least Square**

Sometimes not all data observed have equally quality. Weighted method is used to de-emphasize the data with worse quality and emphasized the one with better quality. The cost function is updated to

$$J(\boldsymbol{\theta}) = \sum_{n=0}^{N-1} w_n \left(x[n] - s[n;\boldsymbol{\theta}] \right)^2 = \left[\mathbf{x} - \mathbf{s}(\boldsymbol{\theta}) \right]^T \mathbf{W} \left[\mathbf{x} - \mathbf{s}(\boldsymbol{\theta}) \right] \tag{2-25}$$

where normally $\mathbf{W}$ is related with the data quality as $\mathbf{W} = \mathbf{C}_{\mathbf{w}}^{-1}$. Normally the solution of (2-25) is the value of $\boldsymbol{\theta}$ that satisfies

$$\frac{\partial J(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = -2 \left[\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right]^T \mathbf{C}_{\mathbf{w}}^{-1} \left[\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right] = \mathbf{0} \tag{2-26}$$

- **LSE and MLE for Gaussian Noise Case**

We known for the Gaussian noise case, MLE is the solution for

$$\frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]^T \mathbf{C}_{\mathbf{w}}^{-1}\left[\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right] = \mathbf{0} \tag{2-27}$$

Therefore the solution is the same as the solution for (2-26). So for Gaussian noise case, MLE and LSE have the same solution. That means even if we assume the noise is Gaussian, but in fact it is not Gaussian, at least we get the LSE solution.

- **Linear Least Squares**

Signal has linear model on the parameter vector $\boldsymbol{\theta}$ can be described as

$$\mathbf{s} = \mathbf{H}\boldsymbol{\theta} \tag{2-28}$$

where $\mathbf{H}$ is a known $N \times p$ observation matrix and assumed full rank. The WLSE is found by minimizing

$$J(\boldsymbol{\theta}) = \left(\mathbf{x} - \mathbf{H}\boldsymbol{\theta}\right)^T \mathbf{W}\left(\mathbf{x} - \mathbf{H}\boldsymbol{\theta}\right) \tag{2-29}$$

The solution is

$$\hat{\boldsymbol{\theta}}_{WLS} = \left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\mathbf{x} \tag{2-30}$$

- **Nonlinear Least Square**

For nonlinear signal model, there is no exploit solution for (2-24), We need the iterative method to solve it. There are two most common approaches: Newton-Raphson and Gauss-Newton. The first one applies iteratively repeat on the linearized cost function $J(\boldsymbol{\theta})$ about the current estimated $\hat{\boldsymbol{\theta}}$. The second one instead applies iteratively repeat on the linearized signal model $\mathbf{s}(\boldsymbol{\theta})$ about the current estimated $\hat{\boldsymbol{\theta}}$. The iterative equations are given as [1]:

- **Newton-Raphson**:

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \left[\mathbf{H}_{\hat{\boldsymbol{\theta}}_k}^T\mathbf{H}_{\hat{\boldsymbol{\theta}}_k} - \sum_{n=0}^{N-1}\mathbf{G}_n(\hat{\boldsymbol{\theta}}_k)\left(x[n] - s_{\hat{\boldsymbol{\theta}}_k}[n]\right)\right]^{-1}\mathbf{H}_{\hat{\boldsymbol{\theta}}_k}^T\left(\mathbf{x} - \mathbf{s}(\hat{\boldsymbol{\theta}}_k)\right) \tag{2-31}$$

- **Gauss-Newton:**

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k + \left[\mathbf{H}_{\hat{\boldsymbol{\theta}}_k}^T \mathbf{H}_{\hat{\boldsymbol{\theta}}_k} \right]^{-1} \mathbf{H}_{\hat{\boldsymbol{\theta}}_k}^T \left(\mathbf{x} - \mathbf{s}(\hat{\boldsymbol{\theta}}_k) \right) \tag{2-32}$$

where $\hat{\boldsymbol{\theta}}_k$ is the k^{th} iterative estimate of $\boldsymbol{\theta}$, $\mathbf{H}_{\hat{\boldsymbol{\theta}}_k}$ is called the Jacobin matrix. It is the first order partials of signal with respect to unknown parameter as,

$$\mathbf{H}_{\hat{\boldsymbol{\theta}}_k} = \left. \frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_k} \tag{2-33}$$

$\mathbf{G}_n(\hat{\boldsymbol{\theta}}_k)$ is called Hessian matrix and it is the second order partials as

$$\mathbf{G}_n(\hat{\boldsymbol{\theta}}_k) = \left. \frac{\partial^2 s_{\boldsymbol{\theta}}[n]}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_k} \tag{2-34}$$

Which one to be chosen depends on the signal model, such as how large is the second order derivative and how complicated it is. In this dissertation, we use Gauss-Newton method and the reasons are discussed later.

In this chapter, we introduction what is parameter estimation. To estimate the unknown parameters, first of all, we need a signal model and PDF to describe the statistical characteristics of observed data, then based on the PDF to calculate one of the most important bound to assess the accuracy of a MVU, the CRLB or the FIM. We also introduced two estimation methods: MLE and LSE, the definition and estimation procedures of them. We use these two methods in this dissertation.

3 TDOA/FDOA Localization-Known and New Foundations

Estimating the location of a passive emitter has been a research issue for decades ([6]~[20]). The procedure is that sensors receive signals sent by the emitter, estimate one or more signal parameter(s) (such as arriving time, carrier frequency, arriving angle and phase) and then use these measurements to estimate emitter location. Regardless of what parameter(s) measures, we need to find a signal model that describe the relationship between the measured parameter(s) and the emitter location first, and then use some estimation methods (such as MLE, LSE) to estimate the emitter location.

There are single-platform method and multiple-platforms method. Single-platform measures the signal parameters consequently within some time intervals. Multiple-platforms measure the signal parameters at the same time. The single one has more flexibility since it only needs one platform to do the task. Multiple-platforms can generally provide higher accuracy and can do multi-tasks simultaneously. In this dissertation we choose multi-platforms methods to get higher estimation accuracy and network flexibility.

Among all the measurements, Time-Difference-of Arrival (TDOA) and Frequency-Difference-of Arrival (FDOA) have been shown to enable highly accurate locations. One measurement of either TDOA or FDOA provides a curve (if assume two dimensional, for three dimensional it is a surface) on which the emitter is known to lie. Two (three for three dimensional, here after we all assume two dimension for illustration simplicity.) such measurements will have an intersection, which is the estimated emitter location. The

signal parameter and geometry properties define the shape of the curves. Since the received data always come with noise, the estimated signal parameter(s) come with errors, and the given sensor's geometry data are not exactly correct, therefore the estimated location is related with all measurements. The main contribution of this dissertation is to find the relationship among them and try to use these properties to have network-wised optimization.

3.1 Overview of TDOA/FDOA Location

3.1.1 Doppler Shift and Time Delay

In order to introduce the TDOA/FDOA localization, we need to give a basic idea about what are Doppler shift and time delay and how they affect the received signals. Assume the signal transmitted by the emitter is $f(t)$, since the sensor is moving the distance between the emitter and sensor is also a function of time as $R(t)$, then the radio frequency (RF) signal propagation time is $\tau(t) = R(t)/c$, the signal received at the sensor can be describe as $f_r(t) = f(t-\tau(t))$. The Doppler shift and time delay are all induced from $R(t)$, which is

$$R(t) = R_0 + v \cdot t + \frac{1}{2} a \cdot t^2 + \cdots \tag{3-1}$$

We keep up to the first order, since in small time interval the velocity change is negligible. Then the received real narrow-band signal can be written as

$$f_r(t) = f(t - [R_0 + v \cdot t]/c) = f([1 - v/c]t - R_0/c) \tag{3-2}$$

where c is the speed of light, $[1-v/c]$ is the time scaling and R_0/c is the time delay names as τ_d. The analytic signal [4] model of the transmitted one can be written as

$$\tilde{f}(t) = E(t)e^{j[\omega_c t+\phi(t)]} \tag{3-3}$$

Then the corresponding received signal is

$$\tilde{f}_r(t) = E([1-v/c]t-\tau_d)e^{j\{\omega_c[1-v/c]t-\tau_d+\phi([1-v/c]t-\tau_d)\}} \tag{3-4}$$

Based on the narrow band approximation [4] $E([1-v/c]t) \approx E(t)$ and $\phi([1-v/c]t) \approx \phi(t)$. Therefore the simplified narrowband low-pass equivalent signal model is

$$\tilde{f}_r(t) = e^{j\alpha}e^{-j\omega_d t}\hat{f}(t-\tau_d) \tag{3-5}$$

where $\alpha = -\omega_c\tau_d$ is the constant term, $\omega_d = \omega_c v/c$ is the Doppler shift term, and $\hat{f}(t-\tau_d) = E(t-\tau_d)e^{j\phi(t-\tau_d)}$ is the time delay term. $\tilde{f}_r(t)$ is the signal that actually gets processed digitally, τ_d and ω_d are the time delay and Doppler shift to be dealing with.

To find the time delay, we need to get the time of arrival (TOA) of $f_r(t)$ at receivers. Assume for k^{th} receiver it is t_k, the unknown time of signal transmitted is t_0, and the distance between the sensor and emitter is d_k. The relationship among these parameters is

$$t_k = t_0 + d_k/c \tag{3-6}$$

d_k can be expressed by terms of the unknown emitter location and known sensor position, t_k can be estimated, but t_0 is not easy or even impossible to estimate. So we can not use (3-6) to estimate the unknown location emitter directly, since there is another unknown

item t_0 in it. We can eliminate the unknown t_0 by subtracting one equation from another; for example, subtract one from its previous one as

$$\Delta t_k = t_k - t_{k-1} = (d_k - d_{k-1})/c \tag{3-7}$$

Δt_k is called the time difference of arrival (TDOA) which can be estimated. Then solving two or more of (3-7) can get the estimation of emitter location which in embedded in d_k and d_{k-1}. Following the same discussing, we can use the frequency difference of arrival (FDOA) of two receivers to estimated the emitter location by solving two or more of the equation

$$\Delta \omega_k = \omega_k - \omega_{k-1} = (\dot{d}_k - \dot{d}_{k-1})/c \tag{3-8}$$

where $\dot{d}_k$ is the time differential of the distance. We will discuss the equations in more details later.

The localization consists of solving a sequence of two estimation problems: (i) processing the intercepted signal samples to estimate the TDOA/FDOA between pairs of sensors, and (ii) processing the TDOA/FDOA estimates to estimate the location of the source.

For simplicity we consider only 2-Dimensional ground-based scenario. We wish to find the location of a stationary emitter, denoted by $\mathbf{p}_e \equiv [x_e, y_e]^T$, and given the N_s sensors, whose positions are $\mathbf{s}_i \equiv [x_i, y_i]^T$ and speeds are $\dot{\mathbf{s}}_i \equiv [\dot{x}_i, \dot{y}_i]^T$, for $i = 1, 2, \ldots, N_s$. Figure 3 is an example of 6 sensors system, there are 3 pairs in this system.

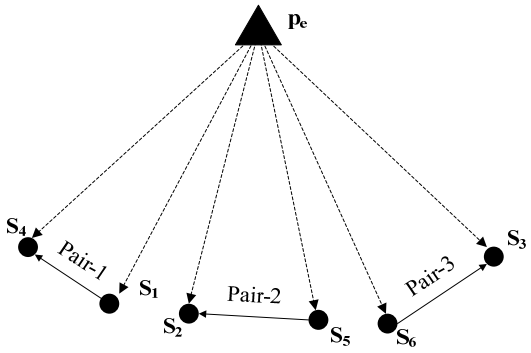

Figure 3 Geometry for stationary source localization system.

3.1.2 First stage: TDOA/FDOA Estimation

The received signal at k^{th} sensor is modeled as

$$x_k(t) = f(t - t_k)e^{j\omega_k t} + w_k(t) \tag{3-9}$$

where $f(t)$ is the transmitted low-pass equivalent signal, t_k is the time delay, ω_k is the Doppler shifts, w_k is the additive white Gaussian noise with zero mean and σ_k^2 as variance, also assumed independent of each other sensors.

Assume the m^{th} pair is paired by i^{th} and j^{th} sensors. The TDOA $\tau_m = t_i - t_j$ and the FDOA $\upsilon_m = \omega_i - \omega_j$ are the parameters to be estimated from time-domain samples of these signals.

From [13], we know the cross-correlation method is used to estimate the TDOA and FDOA between two received signals. The method correlates one received signal with the signal from the paired sensor to find the time delay difference and Doppler shift difference that make the correlation maximum. It is the maximum likelihood estimator. Rewrite the two received signals as

$$\begin{aligned} x_i(t) &= f_i(t) + w_i(t) \\ x_j(t) &= f_i(t-\tau_m)e^{j\upsilon_m t} + w_j(t) \end{aligned} \tag{3-10}$$

where $f_i(t) = f(t-t_i)e^{j\omega_i t}$. Then, use the cross-correlation method on (3-10) to estimate the τ_m and υ_m which maximize the complex ambiguity function as

$$A(\tau,\upsilon) = \int_0^T x_i(t)\overline{x}_j(t+\tau)e^{-j\upsilon t}dt \tag{3-11}$$

Let $\boldsymbol{\theta}_m = [\tau_m, \upsilon_m]^T$ be the parameter vector to be estimated, $\hat{\tau}_m$ and $\hat{\upsilon}_m$ be the estimates, $\Delta\tau_m$ and $\Delta\omega_m$ be the estimation errors, then

$$\begin{aligned} \hat{\tau}_m &= \tau_m + \Delta\tau_m \\ \hat{\upsilon}_m &= \upsilon_m + \Delta\upsilon_m \end{aligned} \tag{3-12}$$

The asymptotic properties of ML estimator gives that the PDF of $\boldsymbol{\theta}_m$ is Gaussian with covariance matrix that is the CRLB of $\boldsymbol{\theta}_m$, so

$$\begin{bmatrix} \Delta\tau_m \\ \Delta\upsilon_m \end{bmatrix} \overset{a}{\sim} N(\mathbf{0}, \mathbf{F}_m^{-1}) \tag{3-13}$$

where $\mathbf{F}_m$ is the FIM of $\boldsymbol{\theta}_m$. Stein's paper [13] gives how to compute the FI for TDOA and FDOA, and we will discuss it in more detail later.

Assume there are *M* pairs in the network. FIM of $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, \cdots, \boldsymbol{\theta}_M^T]^T$ will have the block structure as

$$\mathbf{F}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{F}_1 & \mathbf{CI}_{12} & \cdots & \mathbf{CI}_{1M} \\ \mathbf{CI}_{21} & \mathbf{F}_2 & \ddots & \mathbf{CI}_{2M} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{CI}_{M1} & \mathbf{CI}_{M2} & \cdots & \mathbf{F}_M \end{bmatrix} \tag{3-14}$$

where $\mathbf{CI}_{m,n}$ is the cross term FIM between m^{th} and n^{th} pairs, which is evaluated at session 3.3.

3.1.3 Second Stage: Emitter Location Estimation

In the first stage, we estimate TDOA/FDOA of some pairs of received signals. In the second stage, use these parameters to estimate the emitter location.

- **Signal Model**

Let $\mathbf{r}_k$ be the vector pointing from the emitter to the k^{th} sensor S_k, r_k be the Euclidean distance between them as

$$r_k = \|\mathbf{r}_k\| = \|\mathbf{s}_k - \mathbf{p}_e\| = \sqrt{(x_k - x_e)^2 + (y_k - y_e)^2} \tag{3-15}$$

$\mathbf{u}_k$ be the unit vector of $\mathbf{r}_k$ as $\mathbf{u}_k = \mathbf{r}_k / r_k$, f_e be the transmitted frequency of the transmitter, and c be the speed of light. The vectors used in TDOA/FDOA signal models are illustrated in Figure 4.

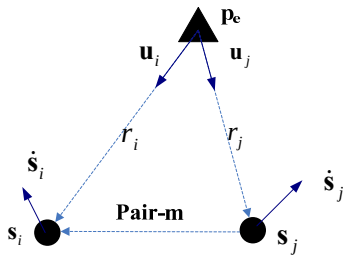

Figure 4 Vectors used in TDOA/FDOA equations illustration for one pair

The signal model of τ_m and υ_m are given by

$$\begin{aligned} \tau_m &= f_m^{\tau}(\mathbf{p}_e) = \frac{1}{c}\left(r_i - r_j\right) \\ \upsilon_m &= f_m^{\upsilon}(\mathbf{p}_e) = \frac{f_e}{c}\left(\mathbf{u}_i^T \dot{\mathbf{s}}_i - \mathbf{u}_j^T \dot{\mathbf{s}}_j\right) \end{aligned} \tag{3-16}$$

The measurement vector $\hat{\mathbf{m}}$ in this stage is the estimation TDAO/FDOA vector of the first stage. Assume *M* pairs in the network. The vector form signal model is

$$\hat{\mathbf{m}} = \mathbf{f}(\mathbf{p}_e) + \mathbf{e} \tag{3-17}$$

where

$$\begin{aligned} \hat{\mathbf{m}} &= \begin{bmatrix} \hat{\tau}_1 & \hat{\upsilon}_1 & \cdots & \hat{\tau}_M & \hat{\upsilon}_M \end{bmatrix}^T \\ \mathbf{e} &= \begin{bmatrix} \Delta\tau_1 & \Delta\upsilon_1 & \cdots & \Delta\tau_M & \Delta\upsilon_M \end{bmatrix}^T \\ \mathbf{f}(\mathbf{p}_e) &= \begin{bmatrix} f_1^{\tau}(\mathbf{p}_e) & f_1^{\upsilon}(\mathbf{p}_e) & \cdots & f_M^{\tau}(\mathbf{p}_e) & f_M^{\upsilon}(\mathbf{p}_e) \end{bmatrix}^T \end{aligned} \tag{3-18}$$

$f_m^{\tau}(\mathbf{p}_e)$ and $f_m^{\upsilon}(\mathbf{p}_e)$ are defined in (3-16). Noise vector $\mathbf{e}$ has the Gaussian distribution with covariance matrix $\mathbf{C}_{\mathbf{e}} = \mathbf{F}^{-1}(\boldsymbol{\theta})$, where $\mathbf{F}(\boldsymbol{\theta})$ is defined as (3-14).

- **CRLB of $\mathbf{p}_e$**

Therefore (3-17) is a deterministic signal plus Gaussian noise model, from (2-20) we get

$$\mathrm{CRLB}(\mathbf{p}_e) = \left[\frac{\partial \mathbf{f}^T(\mathbf{p}_e)}{\partial \mathbf{p}_e} \mathbf{F}(\boldsymbol{\theta}) \frac{\partial \mathbf{f}(\mathbf{p}_e)}{\partial \mathbf{p}_e} \right]^{-1} \tag{3-19}$$

We call $\mathbf{H} = \dfrac{\partial \mathbf{f}(\mathbf{p}_e)}{\partial \mathbf{p}_e}$ Jacobin matrix and evaluated as

$$\mathbf{H} = \frac{\partial \mathbf{f}(\mathbf{p}_e)}{\partial \mathbf{p}_e} = \begin{bmatrix} \dfrac{\partial f_1^{\tau}(\mathbf{p}_e)}{\partial x_e} & \dfrac{\partial f_1^{\tau}(\mathbf{p}_e)}{\partial y_e} \\ \dfrac{\partial f_1^{\upsilon}(\mathbf{p}_e)}{\partial x_e} & \dfrac{\partial f_1^{\upsilon}(\mathbf{p}_e)}{\partial y_e} \\ \vdots & \vdots \\ \dfrac{\partial f_M^{\tau}(\mathbf{p}_e)}{\partial x_e} & \dfrac{\partial f_M^{\tau}(\mathbf{p}_e)}{\partial y_e} \\ \dfrac{\partial f_M^{\upsilon}(\mathbf{p}_e)}{\partial x_e} & \dfrac{\partial f_M^{\upsilon}(\mathbf{p}_e)}{\partial y_e} \end{bmatrix} = \begin{bmatrix} \mathbf{G}_1 \\ \vdots \\ \mathbf{G}_M \end{bmatrix} \tag{3-20}$$

where $\mathbf{G}_m$ is the Jacobin matrix of the m^{th} pair of sensors, defined by $\mathbf{G}_m = \dfrac{\partial \mathbf{f}_m(\mathbf{p}_e)}{\partial \mathbf{p}_e}$ and calculated by

$$\mathbf{G}_m = \begin{bmatrix} \mathbf{u}_i^T - \mathbf{u}_j^T \\ \dfrac{(\mathbf{u}_i^T \dot{\mathbf{s}}_i \mathbf{u}_i - \dot{\mathbf{s}}_i)^T}{r_i} - \dfrac{(\mathbf{u}_j^T \dot{\mathbf{s}}_j \mathbf{u}_j - \dot{\mathbf{s}}_j)^T}{r_j} \end{bmatrix} \tag{3-21}$$

Rewrite the FIM of $\mathbf{p}_e$ as

$$\mathbf{J}_{geo}(\mathbf{p}_e) = \mathbf{H}^T \mathbf{F}(\boldsymbol{\theta})\mathbf{H} \tag{3-22}$$

- **Estimation Method**

We select Weighted Lease Square (WLS) method to estimate $\mathbf{p}_e$, which is to solve

$$\min_{\hat{\mathbf{p}}_e} \left\{ [\mathbf{m} - \mathbf{f}(\hat{\mathbf{p}}_e)]^T \mathbf{W} [\mathbf{m} - \mathbf{f}(\hat{\mathbf{p}}_e)] \right\} \tag{3-23}$$

where $\mathbf{W} = \mathbf{C}_\mathbf{e}^{-1}$ is the weighted matrix. It is obvious that the function vector $\mathbf{f}(\mathbf{p}_e)$ is not linear with respect to the emitter location vector $\mathbf{p}_e$. We need the iterative method (2-31) or (2-32) to solve (3-23). We chose Gauss-Newton method based on: (i) Hessian matrix $\mathbf{G}_n$ is small enough to be negligible; (ii) Error term is small enough to make the sum term negligible; (iii) Inclusion of the sum term can sometimes de-stabilize the iteration; and (iv) For deducing the complexity of computation.

3.2 Importance of Signal Model

In chapter 2, we discussed that the parameter estimation procedure is to find an optimal algorithm which satisfies some rules, such as MMSE or MVU. It is statistical optimal, we need the statistical characteristics of the measured data. The implementation of optimal processing for the second stage of TDOA/FDOA localization requires an understanding of the probabilistic characteristics of the TDOA/FDOA estimates from the first stage. From (3-22), the evaluation of FIM of $\mathbf{p}_e$ needs the calculation of the FIM of

TDOA/FDOA, which is related with signal model of the emitter. In this session, we discuss how the signal model will affect the statistical characteristics and the optimal method. And the conclusion is one of the most important contributions of this dissertation.

Much work has been done to derive optimal TDOA/FDOA estimation methods and to characterize the covariance matrix of the TDOA/FDOA estimates([6]~[20]). TDOA/FDOA results were first developed in the early 1970s for the case of passively locating underwater acoustic sources, where the accepted model for the signal is a stationary random process, which is almost always assumed Gaussian. Then TDOA/FDOA results were developed for the case of passively locating electromagnetic sources, such as radar and communication transmitters. But for the electromagnetic sources, most were deemed still as stationary random process. In fact, a deterministic signal model may be better to apply.

3.2.1 Common and Uncommon Aspects of Acoustic and Electromagnetic Signals

- **Common Aspects for Both Acoustic and Electromagnetic Signals**

The model for two sampled passively-received baseband signals at two sensors is given by

$$\begin{aligned} x_1[n] &= s(nT-\tau_1)e^{jv_1 nT} + w_1[n] \\ x_2[n] &= s(nT-\tau_2)e^{jv_2 nT} + w_2[n] \end{aligned} \tag{3-24}$$

where $s(t)$ is the complex envelope of the transmitted signal, τ_1 and τ_2 are delays and υ_1 and υ_2 are Doppler shifts. The TDOA $\tau_{12} = \tau_1 - \tau_2$ and the FDOA $\upsilon_{12} = \upsilon_1 - \upsilon_2$ are the

parameters to be estimated from time-domain samples of these signals; we define $\boldsymbol{\theta} = \begin{bmatrix} \tau_{12} & \upsilon_{12} \end{bmatrix}^T$. For both the acoustic scenario and the electromagnetic scenario, the accepted modeling assumptions for the noises for $w_1[n]$ and $w_2[n]$ are: (i) They are zero-mean stationary random processes; (ii) They are each Gaussian; (iii) They are independent of each other; and (iv) They are white for simplicity of our discussion. For notational purposes we define: the signal $s_i[n] = s(nT - \tau_i)e^{jv_i nT}$, $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1^T & \mathbf{x}_2^T \end{bmatrix}^T$ and $\mathbf{s}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{s}_2^T \end{bmatrix}^T$, where we explicitly show the dependence on the TDOA/FDOA parameter vector. This much is common between the acoustic and electromagnetic scenarios. The differences arise in what is assumed about the model for the signal $s_i[n]$.

- **Acoustic Signal Model**

For the acoustic scenario the accepted modeling assumptions on $s_i[n]$ are: (i) It is a zero-mean stationary random processes; (ii) It is Gaussian; (iii) It is independent of each noise process; and (iv) It need not be assumed white, although that is a common special case that is considered. For this scenario: (i) The random process assumption is consistent with the erratic nature; (ii) The Gaussian assumption is motivated by the central limit theorem and makes the problem easily tractable; and (iii) The independence from the noises is reasonable based on physical considerations.

- **Electromagnetic Signal Model**

Signals emitted by electromagnetic sources tend to have much more regular structure than the erratic variations seen in acoustic signals made by ocean vehicles and therefore do not readily evoke the notion of random process. A classic example of a stationary random process is a sinusoidal signal with uniformly distributed phase; despite the fact

that each realization of this process exhibits very regular structure it is a random process and it is stationary. Similarly, radar pulse trains can be viewed as random processes for the very same reason: they can be modeled as having random transmission parameters (e.g., random time offset, random phase offset, etc.). However, such signals – with their widely spaced pulses – can hardly be thought to be stationary processes (e.g., variance within a pulse is not equal to the variance between pulses). Furthermore, they certainly cannot be modeled as Gaussian. In such a scenario it is best to consider the signal $s_i[n]$ to be a deterministic signal rather than a random signal.

There is an immediate fundamental distinction between these two models. When we consider the estimates of TDOA and FDOA, $\hat{\tau}_{12}$ and $\hat{\upsilon}_{12}$, we typically wish to find unbiased estimates that minimize $\mathrm{E}\left[\left(\hat{\tau}_{12}-\tau_{12}\right)^2\right]$ and $\mathrm{E}\left[\left(\hat{\upsilon}_{12}-\upsilon_{12}\right)^2\right]$, where it needs to be understood that when the signal is random these expectations are taken over the combined ensemble of signal and noise, such as the Bayesian MSE. When the signal is deterministic these expectations are taken over only the noise ensemble. Thus, when the signal is random we are essentially finding the average squared error over all possible noises and signals; when the signal is deterministic we are essentially finding the average squared error over all possible noises for one specific signal.

3.2.2 Fisher Information for the Two Scenarios

From the above discussion we see that for both the acoustic scenario's model and the electromagnetic scenario's model the received data vector $\mathbf{x}$ is Gaussian and has a Gaussian PDF. The key distinction between these two scenarios that drives all the differences in FIM, CRLB and optimum processing is the manner in which the TDOA

and FDOA impact the parameters of the Gaussian PDF of data vector $\mathbf{x}$. For the case of the acoustic scenario, the mean of $\mathbf{x}$ is zero and the covariance matrix of $\mathbf{x}$ has elements drawn from the autocorrelation function; from this we see that the covariance matrix of $\mathbf{x}$ depends on TDOA/FDOA, so we denote it as $\mathbf{C}(\boldsymbol{\theta})$ to show the dependence. In contrast, for the case of the electromagnetic scenario, the mean of $\mathbf{x}$ is $\mathbf{s}(\boldsymbol{\theta})$ which depends on TDOA/FDOA and the covariance matrix $\mathbf{C}$ of $\mathbf{x}$ is a block diagonal matrix of the two individual noise covariance matrices, and therefore does not depend on TDOA/FDOA. From this single distinction we see that the PDF for the acoustic case is

$$p_{ac}(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}\det^{1/2}[\mathbf{C}(\boldsymbol{\theta})]}\exp\left[-\frac{1}{2}\mathbf{x}^T\mathbf{C}^{-1}(\boldsymbol{\theta})\mathbf{x}\right] \tag{3-25}$$

and the PDF for the electromagnetic case is

$$p_{em}(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}\det^{1/2}[\mathbf{C}]}\exp\left\{-\frac{1}{2}\left[\mathbf{x}-\mathbf{s}(\boldsymbol{\theta})\right]^T\mathbf{C}^{-1}\left[\mathbf{x}-\mathbf{s}(\boldsymbol{\theta})\right]\right\} \tag{3-26}$$

It is this difference that leads to significant differences in the structures of FIM as well as the maximum likelihood estimators for the two cases. For the acoustic scenario the mean of $\mathbf{x}$ is zero and therefore does not depend on the parameter vector; thus the first term in (2-20) is zero and the FIM of TDOA/FDOA $\mathbf{J}_{ac}$ for the acoustic scenario is then given by

$$\left[\mathbf{J}_{ac}(\boldsymbol{\theta})\right]_{ij} = \frac{1}{2}tr\left[\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_i}\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial\mathbf{C}(\boldsymbol{\theta})}{\partial\theta_j}\right] \tag{3-27}$$

However, for the electromagnetic scenario the covariance of $\mathbf{x}$ does not depend on the parameter vector; thus, the second term in (2-20) is zero and the FIM $\mathbf{J}_{em}$ for the electromagnetic scenario is thus given by

$$\left[\mathbf{J}_{em}(\boldsymbol{\theta})\right]_{ij} = \left[\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_i}\right]^T \mathbf{C} \left[\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_i}\right] \tag{3-28}$$

Comparing (3-27) and (3-28) we see that there is a significant difference between the structures of the FIM for the two cases. Furthermore, as pointed out in several publications the off-diagonal elements of the acoustic scenario FIM in (3-27) are zero under a mild assumption of large time-bandwidth product, thus indicating that for the acoustic scenario the optimal estimate of TDOA should be uncorrelated with the optimal estimate of FDOA. However, the electromagnetic scenario FIM in (3-28) does not, in general, yield this uncorrelated TDOA/FDOA condition. For example, for the case of white noise the result in (3-28) gives the cross term Fisher Information element as[25]

$$\left[\mathbf{J}_{em}(\boldsymbol{\theta})\right]_{12} = \mathrm{Re}\left\{\frac{1}{\sigma_1^2}\sum_n -jnTs^*(nT-\tau_1)s'(nT-\tau_1) + \frac{1}{\sigma_2^2}\sum_n -jnTs^*(nT-\tau_2)s'(nT-\tau_2)\right\} \tag{3-29}$$

which in general is not zero. Thus, for the acoustic case we can expect that for a pair of sensors the optimal TDOA estimate is uncorrelated from the optimal FDOA estimate but that should not be expected in the electromagnetic case.

An important impact of this comes when assessing the location accuracy that can be achieved from a set of TDOA/FDOA measurements obtained from M pairs of sensors. Assuming an ML estimator for the TDOA/FDOA values, their estimates can be taken to be Gaussian and then the CRLB on the location estimate covariance becomes (3-22), where $\mathbf{F}(\boldsymbol{\theta})$ is the FIM for all M TDOA/FDOA measurements. When performing studies of location accuracy, we should use the correct FIM $\mathbf{F}(\boldsymbol{\theta})$ for TDOA/FDOA.

3.2.3 Maximum Likelihood Estimator for the Two Scenarios

The MLE for a vector parameter $\boldsymbol{\theta}$ is defined to be the value that maximize the likelihood function $p(\mathbf{x};\boldsymbol{\theta})$ over the allowable domain for $\boldsymbol{\theta}$. For the generalized Gaussian case, it is given by[1]

$$\frac{\partial \ln\{p_{gg}(\mathbf{x};\boldsymbol{\theta})\}}{\partial \theta_i} = -\mathrm{tr}\left(\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i}\right) + \left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_i}\right]^T \mathbf{C}^{-1}(\boldsymbol{\theta})[\mathbf{x}-\boldsymbol{\mu}(\boldsymbol{\theta})] - [\mathbf{x}-\boldsymbol{\mu}(\boldsymbol{\theta})]^T\left[\frac{\partial \mathbf{C}^{-1}(\boldsymbol{\theta})}{\partial \theta_i}\right][\mathbf{x}-\boldsymbol{\mu}(\boldsymbol{\theta})] \quad (3\text{-}30)$$

Notice that there are three terms in this result: one that depends on the sensitivity of the mean to the parameters and two that depend on the sensitivity of the covariance to the parameters. Because the acoustic and electromagnetic scenarios are two different special cases of the generalized Gaussian scenario, we can use the result in (3-30) to find the result for each of these two special cases.

For the acoustic scenario the mean of $\mathbf{x}$ is zero and therefore does not depend on the parameter vector; thus, the partial derivatives of the log-likelihood function (LLF) are given just two terms from (3-30) as

$$\frac{\partial \ln\{p_{ac}(\mathbf{x};\boldsymbol{\theta})\}}{\partial \theta_i} = -\mathrm{tr}\left(\mathbf{C}^{-1}(\boldsymbol{\theta})\frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i}\right) - \mathbf{x}^T\left[\frac{\partial \mathbf{C}^{-1}(\boldsymbol{\theta})}{\partial \theta_i}\right]\mathbf{x} \quad (3\text{-}31)$$

This result is well known in the passive sonar literature. It is stated that for the TDOA/FDOA case the determinant term in the first line of (3-31) does not depend on the parameter vector; therefore, the first term in the second line of (3-31) can be ignored to give

$$\frac{\partial \ln\{p_{ac}(\mathbf{x};\boldsymbol{\theta})\}}{\partial \theta_i} = -\mathbf{x}^T\left[\frac{\partial \mathbf{C}^{-1}(\boldsymbol{\theta})}{\partial \theta_i}\right]\mathbf{x} \quad (3\text{-}32)$$

For the electromagnetic scenario the covariance of $\mathbf{x}$ does not depend on the parameter vector; thus, the partial derivatives of the LLF are given by the second term in (3-30)

$$\frac{\partial \ln\{p_{em}(\mathbf{x};\boldsymbol{\theta})\}}{\partial \theta_i} = \left[\frac{\partial \mathbf{s}(\boldsymbol{\theta})}{\partial \theta_i}\right]^T \mathbf{C}^{-1}[\mathbf{x}-\mathbf{s}(\boldsymbol{\theta})] \tag{3-33}$$

Compare (3-32) and (3-33) shows that we should expect fundamental differences between the MLE for the acoustic and electromagnetic cases. Surprisingly though, each case results in a structure that involves pre-filtering the received signals followed by cross-correlation. However, although both cases share this generalized correlator structure, the pre-filtering needed for each case is quite different. For the acoustic case the filters depend on interplay between the signal power spectra density (PSD) and the noise PSD, whereas for the electromagnetic case the filters depend only on the noise PSD and not on the signal structure.

The underlying assumption about the signal models (i.e., WSS Gaussian signal for the passive acoustic case and a deterministic signal for the passive electromagnetic case) leads to important differences in the results for the FIM, CRB, and MLE. The main differences are that: (i) The general structures of the FIM and CRB are significantly different; (ii) A key specific difference in the FIM/CRB structure is that unlike in the acoustic case, for the electromagnetic case the FDOA and TDOA estimates of a signal pair are likely to be correlated; (iii) For the electromagnetic case the MLE is an unfiltered cross-correlator whenever the noise is white (the acoustic case requires the signal to be white in order to remove the filters). Ignoring these differences can lead to incorrect

location accuracy assessments as well as improper choices when developing processing schemes.

3.3 Characterizing TDOA/FDOA Performance for RF Emitters

In chapter 2, we introduced the characteristics of an estimator and in the previous session, we discussed the importance of signal model when evaluating an estimator. We made some importance conclusions about electromagnetic signal used in the RF TDOA/FDOA system. In this session, we use these conclusions to characterize the TDOA/FDOA performance.

3.3.1 Evaluating the FIM of One Pair

Assume received signals

$$\begin{aligned} x_1[n] &= s[nT-\tau_1]e^{j\nu_1 nT} + \omega_1[n] \\ x_2[n] &= s[nT-\tau_2]e^{j\nu_2 nT} + \omega_2[n] \end{aligned} \tag{3-34}$$

where $s[nT]$ is the sampled transmitted signal, and $\omega_i[n], i=1,2$ is the AWGN received by sensor and assume independent of each other. Vector term of (3-34) is

$$\begin{aligned} \mathbf{x}_i &= \mathbf{s}(\tau_i,\nu_i) + \mathbf{w}_i \\ \mathbf{w}_i &\sim N(\mathbf{0},\mathbf{\Sigma}_i) \end{aligned} \tag{3-35}$$

Let

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} \mathbf{s}_1(\boldsymbol{\theta}) \\ \mathbf{s}_2(\boldsymbol{\theta}) \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \tau_{12} \\ \upsilon_{12} \end{bmatrix} \tag{3-36}$$

where $\tau_{12} = \tau_1 - \tau_2; \upsilon_{12} = \upsilon_1 - \upsilon_2$

Then $\mathbf{x} \sim N(\mathbf{s}(\boldsymbol{\theta}), \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a block diagonal matrix of the two individual noise

covariance matrices such as $\boldsymbol{\Sigma} = diag(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2)$.

Since we assume the RF signal is deterministic, according to (3-28)

$$\left[\mathbf{F}(\boldsymbol{\theta})\right]_{11} = \left(\frac{\partial \mathbf{s}}{\partial \tau_{1,2}}\right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \mathbf{s}}{\partial \tau_{1,2}}\right) \tag{3-37}$$

Since

$$\frac{\partial \mathbf{s}}{\partial \tau_{1,2}} = \begin{bmatrix} \frac{\partial \mathbf{s}_1}{\partial \tau_{1,2}} \\ \frac{\partial \mathbf{s}_2}{\partial \tau_{1,2}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{s}_1}{\partial \tau_1} \cdot \frac{\partial \tau_1}{\partial \tau_{1,2}} \\ \frac{\partial \mathbf{s}_2}{\partial \tau_2} \cdot \frac{\partial \tau_2}{\partial \tau_{1,2}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{s}_1}{\partial \tau_1} \\ -\frac{\partial \mathbf{s}_2}{\partial \tau_2} \end{bmatrix} \tag{3-38}$$

Substitute (3-38) into (3-37), we get

$$\left[\mathbf{F}(\boldsymbol{\theta}_1)\right]_{11} = \left(\frac{\partial \mathbf{s}_1}{\partial \tau_1}\right)^T \boldsymbol{\Sigma}_1^{-1} \left(\frac{\partial \mathbf{s}_1}{\partial \tau_1}\right) + \left(\frac{\partial \mathbf{s}_2}{\partial \tau_2}\right)^T \boldsymbol{\Sigma}_2^{-1} \left(\frac{\partial \mathbf{s}_2}{\partial \tau_2}\right) = F(\tau_1) + F(\tau_2) \tag{3-39}$$

where $F(\tau_i)$ is the Fisher Information of the Time of Arrive (TOA) of the i^{th} sensor, which is evaluated as [13]

$$\begin{aligned} F(\tau) &= 8\pi^2 N \times SNR \times B_{rms}^2 \\ B_{rms}^2 &= \frac{1}{N} \frac{\sum_{k=-N/2}^{N/2-1} k^2 \left|S[k]\right|^2}{\sum_{k=-N/2}^{N/2-1} \left|S[k]\right|^2} \end{aligned} \tag{3-40}$$

where N is the number of samples, $S[k]$ is the DFT value. Following the same rule we get

$$\left[\mathbf{F}(\boldsymbol{\theta})\right]_{12} = \left(\frac{\partial \mathbf{s}}{\partial \tau_{1,2}}\right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \mathbf{s}}{\partial \upsilon_{1,2}}\right) = F(\tau_1, \upsilon_1) + F(\tau_2, \upsilon_2) \tag{3-41}$$

and

$$\left[\mathbf{F}(\boldsymbol{\theta}_1)\right]_{22} = \left(\frac{\partial \mathbf{s}}{\partial \upsilon_{1,2}}\right)^H \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \mathbf{s}}{\partial \upsilon_{1,2}}\right) = F(\upsilon_1) + F(\upsilon_2) \tag{3-42}$$

where $F(\upsilon_i)$ is the Fisher Information of the Doppler of Arrive (DOA) of the i^{th} sensor, which is evaluated as [13]

$$F(\upsilon) = 8\pi^2 N \times SNR \times T_{rms}^2$$
$$T_{rms}^2 = \frac{1}{N} \frac{\sum_{n=-N/2}^{N/2-1} n^2 \left|s[n]\right|^2}{\sum_{n=-N/2}^{N/2-1} \left|s[n]\right|^2} \tag{3-43}$$

And $F(\tau_i, \upsilon_i)$ is the cross term Fisher Information which could be evaluated as (3-29).

From the above, it is clear that FIM of $\boldsymbol{\theta}$ is the summation of FIM of TOA/DOA of each sensor of the pair as

$$\mathbf{F}(\boldsymbol{\theta}) = \mathbf{F}([\tau_1, \upsilon_1]^T) + \mathbf{F}([\tau_2, \upsilon_2]^T) = \mathbf{F}_1 + \mathbf{F}_2 \tag{3-44}$$

3.3.2 Evaluating the FIM of Two Pairs Sharing One Sensor

Three sensors make two pairs as in Figure 5. The vector form of received signal samples is as (3-35) for $i = 1, 2, 3$.

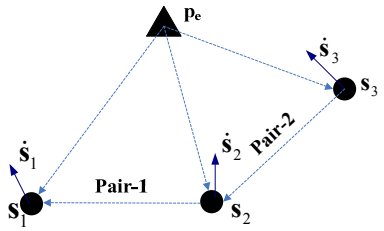

Figure 5 Three sensors and two pairs

Let

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}, \quad \mathbf{s} = \begin{bmatrix} \mathbf{s}_1(\boldsymbol{\theta}) \\ \mathbf{s}_2(\boldsymbol{\theta}) \\ \mathbf{s}_3(\boldsymbol{\theta}) \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \boldsymbol{\theta}_1 \\ \boldsymbol{\theta}_2 \end{bmatrix}, \quad \boldsymbol{\theta}_1 = \begin{bmatrix} \tau_{12} \\ \upsilon_{12} \end{bmatrix} \quad and \quad \boldsymbol{\theta}_2 = \begin{bmatrix} \tau_{23} \\ \upsilon_{23} \end{bmatrix} \tag{3-45}$$

where $\tau_{23} = \tau_2 - \tau_3; \upsilon_{23} = \upsilon_2 - \upsilon_3$

Then $\mathbf{x} \sim N(\mathbf{s}(\boldsymbol{\theta}), \boldsymbol{\Sigma})$, where $\boldsymbol{\Sigma}$ is a block diagonal matrix of the three individual noise covariance matrices such as $\boldsymbol{\Sigma} = diag(\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2, \boldsymbol{\Sigma}_3)$. FIM of $\boldsymbol{\theta}$ will have the following structure

$$\mathbf{J}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{F}(\boldsymbol{\theta}_1) & \mathbf{CI}(\boldsymbol{\theta}) \\ \mathbf{CI}(\boldsymbol{\theta}) & \mathbf{F}(\boldsymbol{\theta}_2) \end{bmatrix} \tag{3-46}$$

According to (3-28) the cross term FIM between pair-1 and pair-2 can be evaluated as

$$\left[\mathbf{CI}(\boldsymbol{\theta})\right]_{11} = \left(\frac{\partial \mathbf{s}}{\partial \tau_{1,2}}\right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \mathbf{s}}{\partial \tau_{1,3}}\right) \tag{3-47}$$

Since

$$\frac{\partial \mathbf{s}}{\partial \tau_{1,2}} = \begin{bmatrix} \frac{\partial \mathbf{s}_1}{\partial \tau_{1,2}} \\ \frac{\partial \mathbf{s}_2}{\partial \tau_{1,2}} \\ \frac{\partial \mathbf{s}_3}{\partial \tau_{1,2}} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{s}_1}{\partial \tau_1} \\ \frac{\partial \mathbf{s}_2}{\partial \tau_2} \\ \mathbf{0} \end{bmatrix} \tag{3-48}$$

and

$$\frac{\partial \mathbf{s}}{\partial \tau_{2,3}} = \begin{bmatrix} \frac{\partial \mathbf{s}_1}{\partial \tau_{2,3}} \\ \frac{\partial \mathbf{s}_2}{\partial \tau_{2,3}} \\ \frac{\partial \mathbf{s}_3}{\partial \tau_{2,3}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \frac{\partial \mathbf{s}_2}{\partial \tau_2} \\ \frac{\partial \mathbf{s}_3}{\partial \tau_3} \end{bmatrix} \tag{3-49}$$

Substitute (3-48) and (3-49) into (3-47), we get

$$\left[\mathbf{CI}(\boldsymbol{\theta})\right]_{11} = \left(\frac{\partial \mathbf{s}_1}{\partial \tau_2}\right)^T \boldsymbol{\Sigma}_1^{-1} \left(\frac{\partial \mathbf{s}_1}{\partial \tau_2}\right) = F(\tau_2) \tag{3-50}$$

Following the same rule to evaluate the other items of we get $\mathbf{CI}(\boldsymbol{\theta})$, it is not hard to get

$$\mathbf{CI}(\boldsymbol{\theta}) = \mathbf{F}_2 \tag{3-51}$$

Therefore the cross term FIM between pairs is the same as the FIM of the shared sensor itself. The FIM of $\boldsymbol{\theta}$ is

$$\mathbf{J}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{F}(\boldsymbol{\theta}_1) & \mathbf{CI}(\boldsymbol{\theta}) \\ \mathbf{CI}(\boldsymbol{\theta}) & \mathbf{F}(\boldsymbol{\theta}_2) \end{bmatrix} = \begin{bmatrix} \mathbf{F}_1 + \mathbf{F}_2 & \mathbf{F}_2 \\ \mathbf{F}_2 & \mathbf{F}_2 + \mathbf{F}_3 \end{bmatrix} \tag{3-52}$$

From the above discuss, the FIM of TDOA/Foes could be calculated for the combination of the FIM of each sensors TOA/DOA itself. That is consistent with Dr. Chen's assessment [23].

3.4 Characterizing Location Performance for RF Emitters

In session 3.3, we discussed about how to compute the FIM of TDOA/FDOA. And it is obvious the estimation accuracy is only related with the received signal quality. But the accuracy of emitter location depends on many items, not only the received signal quality but also the geometry properties of sensors and emitter. The second aspect is the motivation of part of this dissertation: to exploit the geometry characteristics of location estimation and based on them to find optimal pairs.

The location estimation is subject to various types of errors. Emitter location is estimated by some measurements, such as times of arrivals, frequencies of arrivals. The general signal model of any measurement is

$$m = f(\mathbf{p}_e, \mathbf{s}, \dot{\mathbf{s}}) \tag{3-53}$$

where $\mathbf{s}$ and $\dot{\mathbf{s}}$ is the vector of participated sensors' positions and velocities. (3-53) defines a surface on which $\mathbf{p}_e$ lies. The differential of (3-53) with respect to $\mathbf{p}_e, \mathbf{s}$ and $\dot{\mathbf{s}}$ gives

$$dm = \left(\frac{\partial f}{\partial \mathbf{p}_e}\right)^T d\mathbf{p}_e + \left(\frac{\partial f}{\partial \mathbf{s}}\right)^T d\mathbf{s} + \left(\frac{\partial f}{\partial \dot{\mathbf{s}}}\right)^T d\dot{\mathbf{s}} \qquad (3\text{-}54)$$

Errors in the measurement m, $\mathbf{s}$ and $\dot{\mathbf{s}}$ will incur an error in the estimation of $\mathbf{p}_e$. Take variance of both sides of (3-54) gives

$$\sigma_m^2 = \left(\frac{\partial f}{\partial \mathbf{p}_e}\right)^T \mathrm{cov}(\mathbf{p}_e)\left(\frac{\partial f}{\partial \mathbf{p}_e}\right) + \left(\frac{\partial f}{\partial \mathbf{s}}\right)^T \mathrm{cov}(\mathbf{s})\left(\frac{\partial f}{\partial \mathbf{s}}\right)^T + \left(\frac{\partial f}{\partial \dot{\mathbf{s}}}\right)^T \mathrm{cov}(\dot{\mathbf{s}})\left(\frac{\partial f}{\partial \dot{\mathbf{s}}}\right) \qquad (3\text{-}55)$$

This gives an overview of how all these measurements errors play in the geo-location estimation accuracy.

For the TDOA/FDOA methods we choose, and assume two dimensional, the intersection of the two curves is a random variable whose statistics is induces by statistics on the measured quantities, such as the accuracy of TDOA/FDOA estimates, the accuracy of sensors locations and velocities. Figure 6 shows the randomness of the intersection of two TDOA hyperbolas. The bold line is the true hyperbola, the "…" line is the measured/estimated hyperbola from the first pair, and "---" is from the second pair. The randomness can be evaluated by the estimation variance. We discuss all the related aspects about the location estimation accuracy in this session.

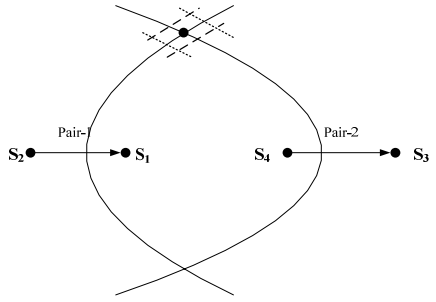

Figure 6 Randomness of the intersection of two TDOA hyperbolas

3.4.1 Emitter Location Accuracy as a Function of TDOAs and Sensor States Measurements

One TDOA value is a function of emitter location and two sensors positions as

$$\tau = f(\mathbf{p}_e, \mathbf{s}_1, \mathbf{s}_2) = \frac{1}{c}(\mathbf{r}_1 - \mathbf{r}_2) \tag{3-56}$$

where $\mathbf{r}_k = \mathbf{p}_e - \mathbf{s}_k$. One this function can define a hyperbola, and two hyperbolas define an intersection. For the sensors as in Figure 6, the two functions are

$$\begin{aligned} \tau_1 &= f_1(\mathbf{p}_e, \mathbf{s}_1, \mathbf{s}_2) \\ \tau_2 &= f_2(\mathbf{p}_e, \mathbf{s}_3, \mathbf{s}_4) \end{aligned} \tag{3-57}$$

Taking differentials of (3-57) and let $\mathbf{q}_1 = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{s}_2^T \end{bmatrix}^T$ and $\mathbf{q}_2 = \begin{bmatrix} \mathbf{s}_3^T & \mathbf{s}_4^T \end{bmatrix}^T$, then

$$\begin{aligned} d\tau_1 &= \left(\frac{\partial f_1}{\partial \mathbf{p}_e}\right)^T d\mathbf{p}_e + \left(\frac{\partial f_1}{\partial \mathbf{q}_1}\right)^T d\mathbf{q}_1 \\ d\tau_2 &= \left(\frac{\partial f_2}{\partial \mathbf{p}_e}\right)^T d\mathbf{p}_e + \left(\frac{\partial f_2}{\partial \mathbf{q}_2}\right)^T d\mathbf{q}_2 \end{aligned} \tag{3-58}$$

Factor $d\mathbf{p}_e$ into two parts as one is in the direction of the normal of the surface defined by f_m and noted as $d\mathbf{p}_e^{\perp}$, which is the direction as $\partial f_m / \partial \mathbf{p}_e$, the other one is perpendicular to the normal vector and noted as $d\mathbf{p}_e^{\parallel}$. The fact oration is illustrated in Figure 7.

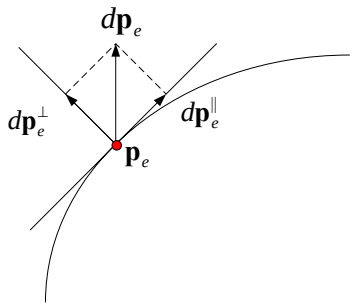

Figure 7 Fact orating $d\mathbf{p}_e$ into two parts

Then $d\mathbf{p}_e$ is factored with f_1 and f_2 respectively as

$$\begin{aligned} d\mathbf{p}_e &= d\mathbf{p}_e^{\perp_1} + d\mathbf{p}_e^{\parallel_1} \\ d\mathbf{p}_e &= d\mathbf{p}_e^{\perp_2} + d\mathbf{p}_e^{\parallel_2} \end{aligned} \tag{3-59}$$

According to vector inner product theory (3-58) becomes

$$\begin{aligned} d\tau_1 &= \left\|\nabla \mathbf{h}_1\right\| \cdot \left\|\Delta \mathbf{n}_1\right\| + \sum_1 \\ d\tau_2 &= \left\|\nabla \mathbf{h}_2\right\| \cdot \left\|\Delta \mathbf{n}_2\right\| + \sum_2 \end{aligned} \tag{3-60}$$

Where for notation simplicity let $\nabla \mathbf{h}_m = \partial f_m / \partial \mathbf{p}_e$, $\Delta \mathbf{n}_m = d\mathbf{p}_e^{\perp_m}$ and

$\sum_m = \left(\partial f_m / \partial \mathbf{q}_m\right)^T d\mathbf{q}_m = \sum_{i=1}^{4} \left(\partial f_m / \partial q_i\right) q_i$, where q_i is corresponding to sensors position parameters.

Assume the measurements of τ_m and $\mathbf{s}_i$ are made. Errors of there measurements will induce an error in the estimate of the intersection where the emitter supposed to be located. Let $\sigma_{n_m}^2$ be emitter location estimation variance on the direction of the normal to the surface of hyperbola (specified by f_m) at emitter location. $\sigma_{\tau_m}^2$ be the TDOA error variance. $\sigma_{q_i}^2$ be the variance of sensors position measurements (since for TDOA only, do not need velocity). And assume they are independent of each other. Rewrite (3-60) as the variance form

$$\sigma_{n_1}^2 = \frac{1}{\left\|\nabla \mathbf{h}_1\right\|^2}\left(\sigma_{\tau_1}^2 + \sigma_{\Sigma_1}^2\right)$$
$$\sigma_{n_2}^2 = \frac{1}{\left\|\nabla \mathbf{h}_2\right\|^2}\left(\sigma_{\tau_2}^2 + \sigma_{\Sigma_2}^2\right) \tag{3-61}$$

Since

$$\left\|\Delta \mathbf{n}_m\right\| = \frac{\nabla \mathbf{h}_m^T}{\left\|\nabla \mathbf{h}_m\right\|} d\mathbf{p}_e \tag{3-62}$$

Then

$$d\mathbf{p}_e = \mathbf{H}^{-1}\Delta \mathbf{n} \tag{3-63}$$

Where

$$\mathbf{H} = \begin{bmatrix} \dfrac{\nabla \mathbf{h}_1^T}{\left\|\nabla \mathbf{h}_1\right\|} \\ \dfrac{\nabla \mathbf{h}_2^T}{\left\|\nabla \mathbf{h}_2\right\|} \end{bmatrix}, \quad \Delta \mathbf{n} = \begin{bmatrix} \left\|\Delta \mathbf{n}_1\right\| \\ \left\|\Delta \mathbf{n}_2\right\| \end{bmatrix} \tag{3-64}$$

Therefore

$$\mathbf{C}_{\mathbf{p}_e} = \mathbf{H}^{-1}\mathbf{C}_{\Delta \mathbf{n}}(\mathbf{H}^{-1})^T \tag{3-65}$$

where

$$\mathbf{C}_{\Delta \mathbf{n}} = \begin{bmatrix} \dfrac{1}{\left\|\nabla \mathbf{h}_1\right\|^2}\left(\sigma_{\tau_1}^2 + \sigma_{\Sigma_1}^2\right) & 0 \\ 0 & \dfrac{1}{\left\|\nabla \mathbf{h}_2\right\|^2}\left(\sigma_{\tau_2}^2 + \sigma_{\Sigma_2}^2\right) \end{bmatrix} \tag{3-66}$$

Computation the two important measurements of $\mathbf{C}_{\mathbf{p}_e}$ are got as

$$\text{trace}(\mathbf{C}_{\mathbf{p}_e}) = \frac{\left(\sigma_{\tau_1}^2 + \sigma_{\Sigma_1}^2\right) + \left(\sigma_{\tau_2}^2 + \sigma_{\Sigma_2}^2\right)}{\left\|\mathbf{H}\right\|^2}$$
$$\text{determinant}(\mathbf{C}_{\mathbf{p}_e}) = \frac{\left(\sigma_{\tau_1}^2 + \sigma_{\Sigma_1}^2\right) \cdot \left(\sigma_{\tau_2}^2 + \sigma_{\Sigma_2}^2\right)}{\left\|\mathbf{H}\right\|^2 \cdot \left\|\nabla \mathbf{h}_1\right\|^2 \cdot \left\|\nabla \mathbf{h}_2\right\|^2} \tag{3-67}$$

Put into the geometrical factors and define a new induced vector concept as

$\mathbf{x}^{\perp} = \begin{bmatrix} \cos(\pi/2) & \sin(\pi/2) \\ \sin(\pi/2) & \cos(\pi/2) \end{bmatrix} \cdot \mathbf{x}$, which means $\mathbf{x}^{\perp}$ is generated by count-clockwise rotating $\mathbf{x}$ by $\pi/2$ degree. Define angles as following

(i) $\theta_1 = \angle(\mathbf{r}_1, \mathbf{r}_2)$ and $\theta_2 = \angle(\mathbf{r}_3, \mathbf{r}_4)$

(ii) $\alpha_{(i,j)} = \angle(\mathbf{r}_i, \mathbf{r}_j^{\perp})$

Then after simplification and computation (3-67) becomes

$$\text{trace}(\mathbf{C}_{\Delta\mathbf{n}}) = \frac{4(1-\cos\theta_1)(1-\cos\theta_2)\left[(\sigma_{\tau_1}^2 + \sigma_{\Sigma_1}^2) + (\sigma_{\tau_2}^2 + \sigma_{\Sigma_2}^2)\right]}{\left(\cos\alpha_{(1,3)} + \cos\alpha_{(2,4)} - \cos\alpha_{(1,4)} - \cos\alpha_{(2,3)}\right)^2}$$

$$\text{determinant}(\mathbf{C}_{\Delta\mathbf{n}}) = \frac{(\sigma_{\tau_1}^2 + \sigma_{\Sigma_1}^2)(\sigma_{\tau_2}^2 + \sigma_{\Sigma_2}^2)}{\left(\cos\alpha_{(1,3)} + \cos\alpha_{(2,4)} - \cos\alpha_{(1,4)} - \cos\alpha_{(2,3)}\right)^2} \quad (3\text{-}68)$$

From (3-68), the variance of estimation errors is a complex combination of both measurement factors (all σ values) and geometrical factors (angles). Based on the relationship of all these angles, $\text{trace}(\mathbf{C}_{\Delta\mathbf{n}})$ may be get more simplified, but even further simpler there is not a clear optimal relationship between the $\text{trace}(\mathbf{C}_{\Delta\mathbf{n}})$ and some specified parameters. It is also for the case of $\text{determinant}(\mathbf{C}_{\Delta\mathbf{n}})$. And there are trade-off among these geometrical factors.

3.4.2 Emitter Location Accuracy as a Function of FDOA and Sensor States Measurements

One FDOA value is a function of emitter location and two sensors positions and velocities as

$$\upsilon = g(\mathbf{p}_e, \mathbf{s}_1, \mathbf{s}_2, \dot{\mathbf{s}}_1, \dot{\mathbf{s}}_2) = \frac{f_e}{c}\left(\frac{\mathbf{r}_1^T \dot{\mathbf{s}}_1}{\|\mathbf{r}_1\|} - \frac{\mathbf{r}_2^T \dot{\mathbf{s}}_2}{\|\mathbf{r}_2\|} \right) \tag{3-69}$$

One this function can define a curve, and two FDOA measurements can define an intersection, for the sensors as in Figure 6, the two functions are

$$\begin{aligned} \upsilon_1 &= g_1(\mathbf{p}_e, \mathbf{s}_1, \mathbf{s}_2, \dot{\mathbf{s}}_1, \dot{\mathbf{s}}_2) \\ \upsilon_2 &= g_2(\mathbf{p}_e, \mathbf{s}_3, \mathbf{s}_4, \dot{\mathbf{s}}_3, \dot{\mathbf{s}}_4) \end{aligned} \tag{3-70}$$

Following the same discussion as TDOA case, we can also get the similar expression as (3-65) with different $\mathbf{H}$ and $\mathbf{C}_{\Delta \mathbf{n}}$ matrices as

$$\begin{aligned} \|\nabla \mathbf{h}_1\| = &\left(\|\dot{\mathbf{s}}_1\|/\|\mathbf{r}_1\|\right)^2 \sin^2 \alpha_1 + \left(\|\dot{\mathbf{s}}_2\|/\|\mathbf{r}_2\|\right)^2 \sin^2 \alpha_2 \\ &- 2\left(\|\dot{\mathbf{s}}_1\|\|\dot{\mathbf{s}}_2\|/\|\mathbf{r}_1\|\|\mathbf{r}_2\|\right)^2 \left(\cos\theta_1 \cos\alpha_1 \cos\alpha_2 - \cos\alpha_1 \cos\beta_1 - \sin\alpha_2 \sin\beta_2\right) \end{aligned} \tag{3-71}$$

The angles are defined as following and illustrated in Figure 8.

(i) $\alpha_1 = \angle(\dot{\mathbf{s}}_1, \mathbf{r}_1)$ and $\alpha_2 = \angle(\dot{\mathbf{s}}_2, \mathbf{r}_2)$

(ii) $\beta_1 = \angle(\dot{\mathbf{s}}_2, \mathbf{r}_1)$ and $\beta_2 = \angle(\dot{\mathbf{s}}_1, \mathbf{r}_2)$

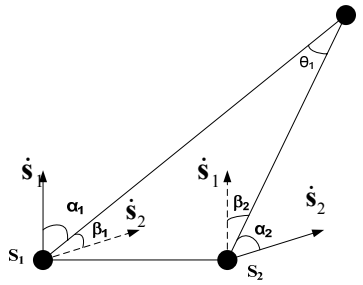

Figure 8 Angles illustration for FDOAs system

The expression of $\|\mathbf{H}\|$ is more complicated with many angles made by sensors' velocities and positions, between sensor itself or cross one. $\mathbf{x}^{\perp}$ is defined in last section. Define angles as following

(i) $\alpha_i = \angle(\mathbf{r}_i, \dot{\mathbf{s}}_i)$

(ii) $\gamma_{(i,j)} = \angle\left(\mathbf{r}_i, \mathbf{r}_j^{\perp}\right)$

(iii) $\lambda_{(i,j)} = \angle\left(\mathbf{r}_i, \dot{\mathbf{s}}_j^{\perp}\right)$

(iv) $\beta_{(i,j)} = \angle\left(\dot{\mathbf{s}}_i, \dot{\mathbf{s}}_j^{\perp}\right)$

Based on these definitions, $\|\mathbf{H}\|$ is given as

$$\|\mathbf{H}\| = \frac{f_{(1,3)} - f_{(1,4)} - f_{(2,3)} + f_{(2,4)}}{\|\nabla \mathbf{h}_1\|^2 \cdot \|\nabla \mathbf{h}_2\|^2} \tag{3-72}$$

where

$$f_{(i,j)} = \frac{\|\dot{\mathbf{s}}_i\|\|\dot{\mathbf{s}}_j\|}{\|\mathbf{r}_i\|\|\mathbf{r}_j\|}\left[\cos\alpha_i \cos\alpha_j \cos\gamma_{(i,j)} - \cos\alpha_i \cos\lambda_{(i,j)} - \cos\alpha_j \cos\lambda_{(j,i)} + \cos\beta_{(i,j)}\right] \tag{3-73}$$

Totally there are sixteen angles. Even though some of them may be represented by some others, (3-72) is still a long expression with many items. Put (3-71) and (3-72) into (3-67) will get two more long expressions, we do not need to give the long expression at here.

The objective of this session is to give an explicit expression about how are all the measurements accuracy and geometrical properties related with the emitter location accuracy. The relationship in general form is given by (3-67). And we checked two special case of (i) two TDOAs and (ii) two FDOAs. The purpose of these derivations is not to give a simple relationship between the accuracy and any related parameter. It is to give an overall insight about what will reflect our estimation accuracy. In the following chapters we are going to discuss more about some specified errors and give out optimal or sub-optimal solutions to solve the unexpected uncertainties.

3.5 Appendix 3A-Convergence of Gauss-Newton Least Square Method

In the second stage of the data processing we use the iterative lease square method to estimate emitter location. Given an initial reference estimation of $\mathbf{p}_e$, iteratively update by (2-32). We need to consider the convergence problem.

Least Square (LS) procedure estimate model parameters $\boldsymbol{\theta}$ by minimizing the LS error criterion, or called cost function

$$J(\boldsymbol{\theta}) = [\mathbf{x} - \mathbf{s}(\boldsymbol{\theta})]^T [\mathbf{x} - \mathbf{s}(\boldsymbol{\theta})] \tag{3-74}$$

In general, $\mathbf{s}(\boldsymbol{\theta})$ is a N-dimensional nonlinear function of $\boldsymbol{\theta}$.

The Newton-Raphson iteration method is to linearize the differentiation of J, which needs the first-order and second-order partials of $\mathbf{s}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. The Gauss-Newton iteration method is to linearize $\mathbf{s}(\boldsymbol{\theta})$ about some nominal $\boldsymbol{\theta}_0$, which only needs the first-order partial of $\mathbf{s}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. The iterative equation is

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + [\mathbf{H}^T(\boldsymbol{\theta}_k)\mathbf{H}(\boldsymbol{\theta}_k)]^{-1}\mathbf{H}(\boldsymbol{\theta}_k)[\mathbf{x} - \mathbf{s}(\boldsymbol{\theta}_k)] \tag{3-75}$$

where

$$[\mathbf{H}(\boldsymbol{\theta})]_{ij} = \frac{\partial s_i(\boldsymbol{\theta})}{\partial \theta_j} \tag{3-76}$$

For small $\|\mathbf{h}\|$ the Taylor expansion of $\mathbf{s}(\boldsymbol{\theta})$ and keep up to the first order derivative item is

$$\mathbf{s}(\boldsymbol{\theta}_k + \mathbf{h}) \approx \mathbf{s}(\boldsymbol{\theta}_k) + \mathbf{H}(\boldsymbol{\theta}_k)\mathbf{h} \tag{3-77}$$

For notation purpose write $\mathbf{s}(\boldsymbol{\theta}_k)$ as $\mathbf{s}_{\boldsymbol{\theta}_k}$ and $\mathbf{H}(\boldsymbol{\theta}_k)$ as $\mathbf{H}_{\boldsymbol{\theta}_k}$. Inserting (3-77) into (3-74) we see that

$$\begin{aligned} J(\boldsymbol{\theta}_k + \mathbf{h}) &= (\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k} + \mathbf{H}_{\boldsymbol{\theta}_k}\mathbf{h})^T (\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k} + \mathbf{H}_{\boldsymbol{\theta}_k}\mathbf{h}) \\ &= (\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k})^T (\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k}) + 2(\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k})^T \mathbf{H}_{\boldsymbol{\theta}_k}\mathbf{h} + \mathbf{h}^T \mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{H}_{\boldsymbol{\theta}_k}\mathbf{h} \end{aligned} \tag{3-78}$$

Let

$$\begin{aligned} L(\mathbf{h}) &= J(\boldsymbol{\theta}_k + \mathbf{h}) \\ \mathbf{y} &= \mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k} \end{aligned} \tag{3-79}$$

then

$$L(\mathbf{h}) = J_{\boldsymbol{\theta}_k} + 2\mathbf{y}^T \mathbf{H}_{\boldsymbol{\theta}_k}\mathbf{h} + \mathbf{h}^T \mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{H}_{\boldsymbol{\theta}_k}\mathbf{h} \tag{3-80}$$

The gradient and the Hessian of $L(\mathbf{h})$ with respect to $\mathbf{h}$ are

$$\begin{aligned} L'(\mathbf{h}) &= 2\mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{y} + 2\mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{H}_{\boldsymbol{\theta}_k}\mathbf{h} \\ L''(\mathbf{h}) &= 2\mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{H}_{\boldsymbol{\theta}_k} \end{aligned} \tag{3-81}$$

In TDOA/FDOA emitter localization system, $\mathbf{H}_{\boldsymbol{\theta}_k}$ can be calculated as (3-20). So generally the columns of $\mathbf{H}_{\boldsymbol{\theta}_k}$ are independent, therefore $rank(\mathbf{H}_{\boldsymbol{\theta}_k}) = 2$, $\mathbf{H}_{\boldsymbol{\theta}_k}$ is full rank. So $L''(\mathbf{h})$ is positive definite. $L(\mathbf{h})$ has a unique minimizer and can be solved as

$$\mathbf{h}^* = -(\mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{H}_{\boldsymbol{\theta}_k})^{-1} \mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{y} \tag{3-82}$$

and

$$J(\boldsymbol{\theta}_k + \mathbf{h}^*) \approx J(\boldsymbol{\theta}_k) + \mathbf{h}^{*T} J'(\boldsymbol{\theta}_k) \tag{3-83}$$

Since the gradient of $J(\boldsymbol{\theta}_k)$ is

$$J'(\boldsymbol{\theta}_k) = \mathbf{H}_{\boldsymbol{\theta}_k}(\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k}) \tag{3-84}$$

then

$$\begin{aligned}\mathbf{h}^{*T} J'(\boldsymbol{\theta}_k) &= \mathbf{h}^{*T} \mathbf{H}_{\boldsymbol{\theta}_k} (\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k}) \\ &= -(\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k})^T (\mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{H}_{\boldsymbol{\theta}_k})^{-1} \mathbf{H}_{\boldsymbol{\theta}_k}^T \mathbf{H}_{\boldsymbol{\theta}_k} (\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k}) \\ &= -(\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k})^T (\mathbf{x} - \mathbf{s}_{\boldsymbol{\theta}_k}) < 0\end{aligned} \tag{3-85}$$

Therefore

$$J(\boldsymbol{\theta}_k + \mathbf{h}) < J(\boldsymbol{\theta}_k) \tag{3-86}$$

So $\mathbf{h}$ is a descent direction of J. So updating $\boldsymbol{\theta}$ by $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \mathbf{h}$, finally the solution can converge to the optimal one.

In this chapter, we (i) Give basic background about emitter location estimation and discuss why we choose TDOA/FDOA methods in this dissertation; (ii) Talk about the basic procedures of the localization; (iii) Exploit the importance of signal model and get the important and fundamental insight about the FIM/CRLB and MLE, which is one of this dissertation's contribution; (iv) Based on the results we get, compute the characteristics of TDOA/FDOA estimation; (v) Develop the geometrical factors importance in the localization, which is the motivation of the next chapter.

4 Network Management---FIM Based Sensor Selection and Pairing

Briefly, the localization procedure is: the full set of sensors is tasked to simultaneously intercept signal data in a specified frequency band and then cooperatively share data to detect the presence of an emitter and then to locate it. After data collection at the sensors, they send a small amount of data to a central node where the data sets are processed to determine which sensors detected data from the same emitter of interest. And based upon this small amount of data, it is possible to determine a rough estimate of the emitter location, which will be used as initial estimation at iterative lease-square estimation.

On the basis of this initial, small amount of shared data from the remaining subset of participating sensors, it is also possible to assess the impact of the relative emitter-sensor geometry on the location processing task, thus allowing subsequent processing to be optimized with respect to the geometry and error sources. Thus, the central node then uses knowledge of the current positions and trajectories of the remaining sensors to further reduce the participating subset based on the quality and the error sensitivity of their data sets. For example, one sensor may have high-quality data (e.g., SNR, Fisher information, etc.) but its position and trajectory give it a high susceptibility to non-noise error sources (e.g., GDOP, sensor vibrations, navigation errors, etc.), another sensor could have low-quality data but have low susceptibility to error sources, etc. By eliminating sensors that have negligible usefulness to the final outcome of the task it is possible to reduce the amount of network communication resources needed to accomplish

the task. Further reduction in the needed communication resources is then achieved through location-optimized compression when this final subset of participating sensors shares its data to support the location-estimation tasks[23]. Then, location processing is completed using algorithms specifically designed to mitigate the impact of the error sources. If at this point the accuracy of the location estimate is desired to be further improved, it is possible to determine optimal trajectories for the selected sensor subset in order to minimize their sensitivities to GDOP, vibration, etc. during subsequent data collection . We briefly discussed this issue in chapter-6. We aim to provide a comprehensive method of maximizing the effectiveness of a sub-set of networked sensors by considering the network-wide setting in which the sensors work. The following sections provide the details upon which these ideas are based.

4.1 Optimal Criterion

We know the Cramer-Rao Lower Bound of a parameter is depended on the signal model. The signal model is related with both the parameters assumed known and the ones to be estimated. Assume a signal model is a function of both $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ as $s(\boldsymbol{\alpha};\boldsymbol{\beta})$, where $\boldsymbol{\alpha}$ is the vector to be estimated and $\boldsymbol{\beta}$ is the system parameter vector which is assumed known and can be modified. The natural and ultimate choice of criteria to optimal an estimation algorithm relies on their error covariance matrix which is lower bounded by the CRLB matrix, which is given in chapter 2. For example, if $\mathrm{CRLB}(\boldsymbol{\alpha};\boldsymbol{\beta}_1) < \mathrm{CRLB}(\boldsymbol{\alpha};\boldsymbol{\beta}_2)$, we can say that the signal model based on $\boldsymbol{\beta}_1$ is better than the one on $\boldsymbol{\beta}_2$ at the estimating of $\boldsymbol{\alpha}$ point of view. For vector case, "<" means

$\mathrm{CRLB}(\boldsymbol{\alpha};\boldsymbol{\beta}_2) - \mathrm{CRLB}(\boldsymbol{\alpha};\boldsymbol{\beta}_1)$ is a positive definite matrix. So if we can choose $\boldsymbol{\beta}$ to get minimal CRLB of $\boldsymbol{\alpha}$, the optimal problem will be given as

$$\boldsymbol{\beta}^* = \arg\min_{\boldsymbol{\beta}} \left\{ \text{some scalar value function of } \mathrm{CRLB}_{\boldsymbol{\alpha}}(\boldsymbol{\alpha};\boldsymbol{\beta}) \right\} \tag{4-1}$$

Since for the vector estimator, the CRLB is a matrix, we need some scalar value function to evaluate the minimal of a matrix, such as eigen-values, trace or determinant.

4.1.1 Covariance Matrix and Error Ellipsoid

The covariance matrix of estimation errors can be visualized in the space of errors. Assume $\hat{\boldsymbol{\theta}}$ is a $p \times 1$ unbiased estimated vector and it is Gaussian with zeros mean and covariance matrix $\mathbf{C}_{\hat{\boldsymbol{\theta}}}$. Then the PDF of $\hat{\boldsymbol{\theta}}$ is given by

$$p(\hat{\boldsymbol{\theta}}) = \frac{1}{(2\pi)^p \left| \mathbf{C}_{\hat{\boldsymbol{\theta}}} \right|^{1/2}} \exp\left\{ -\frac{1}{2} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T \mathbf{C}_{\hat{\boldsymbol{\theta}}}^{-1} (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \right\} \tag{4-2}$$

The equal height contours of this PDF are given by the values of estimation error vector $\Delta\hat{\boldsymbol{\theta}} = \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}$ such that

$$\Delta\hat{\boldsymbol{\theta}}^T \mathbf{C}_{\hat{\boldsymbol{\theta}}}^{-1} \Delta\hat{\boldsymbol{\theta}} = k \tag{4-3}$$

From [17], choosing the value of k, such that

$$k = -2\ln(1 - P_e) \tag{4-4}$$

where P_e is the probability that $\Delta\hat{\boldsymbol{\theta}}$ will lie inside the ellipsoid specified by (4-3). For example, if we want to see the ellipsoid that includes 50% estimates errors, then $k = -2\ln(1 - 0.5)$. The size and orientation of the ellipsoid can be described in terms of eigenvalues and eigenvectors of the symmetric $p \times p$ matrix $\mathbf{C}_{\hat{\boldsymbol{\theta}}}$. Let the eigen-decomposition of $\mathbf{C}_{\hat{\boldsymbol{\theta}}}$ be

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = \sum_{i=1}^{p} \lambda_i \mathbf{v}_i^T \mathbf{v}_i \tag{4-5}$$

where λ_i is the eigenvalue, $\mathbf{v}_i$ is the corresponding eigenvector.

For visualizing, assume $p = 2$ and $\hat{\boldsymbol{\theta}} = [\hat{x}_e, \hat{y}_e]^T$. Therefore the vicariate covariance matrix can be expressed as

$$\mathbf{C}_{\hat{\boldsymbol{\theta}}} = \begin{bmatrix} \sigma_{\hat{x}_e}^2 & \sigma_{\hat{x}_e,\hat{y}_e} \\ \sigma_{\hat{x}_e,\hat{y}_e} & \sigma_{\hat{y}_e}^2 \end{bmatrix} \tag{4-6}$$

and

$$\begin{aligned} \lambda_1 &= \frac{1}{2}\left(\sigma_{\hat{x}_e}^2 + \sigma_{\hat{y}_e}^2 + \sqrt{(\sigma_{\hat{x}_e}^2 - \sigma_{\hat{y}_e}^2)^2 + 4\sigma_{\hat{x}_e,\hat{y}_e}^2}\right) \\ \lambda_2 &= \frac{1}{2}\left(\sigma_{\hat{x}_e}^2 + \sigma_{\hat{y}_e}^2 - \sqrt{(\sigma_{\hat{x}_e}^2 - \sigma_{\hat{y}_e}^2)^2 + 4\sigma_{\hat{x}_e,\hat{y}_e}^2}\right) \end{aligned} \tag{4-7}$$

The error ellipse of $\begin{bmatrix} \Delta\hat{x}_e & \Delta\hat{y}_e \end{bmatrix}^T$ is as in Figure 9. The lengths of the principle axes are proportional to the value of the eigenvalues. And the directions of the principle axes are along with the eigenvectors. It is clear, the smaller λ_i, the more accurate of the estimates, on the contrary, the larger λ_i, the wider estimation error range will be.

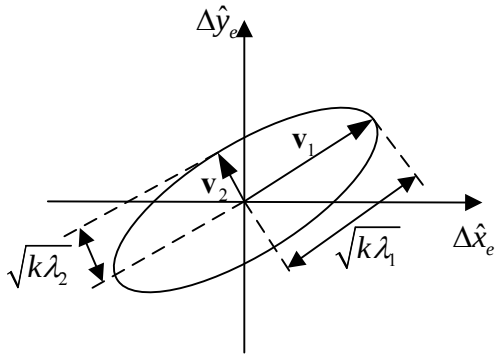

Figure 9 Error ellipse and coordinate axes in the

relationship with eigenvalues and eigenvectors.

4.1.2 Criterion Selection

There are two normally used measurements to measure the size of an ellipse, one is the area and the other one is the perimeter. The calculations of the ellipse in Figure 9 are given by

$$\begin{aligned} \text{Area} &= \pi k\sqrt{\lambda_1\lambda_2} \\ \text{Perimater} &\approx \pi\sqrt{2k}\sqrt{\lambda_1+\lambda_2} \end{aligned} \tag{4-8}$$

From matrix theory, the determinant of a matrix is equal to the product of the eigenvalues, and the trace of a matrix is equal to the summation of the eigenvalues. So the two measurements of the ellipse specified by $\mathbf{C}_{\hat{\boldsymbol{\theta}}}$ are proportional to the two important properties of $\mathbf{C}_{\hat{\boldsymbol{\theta}}}$ as

$$\begin{aligned} \text{Area} &\propto \text{Determinant}(\mathbf{C}_{\hat{\boldsymbol{\theta}}}) \\ \text{Perimeter} &\propto \text{Trace}(\mathbf{C}_{\hat{\boldsymbol{\theta}}}) \end{aligned} \tag{4-9}$$

The inverse of Fisher Information Matrix is the CRLB matrix, which is the "smallest" matrix (as indicated in (2-17)) that an unbiased estimator can get. From chapter 3, the FIM of emitter location is $\mathbf{J}_{geo}(\mathbf{p}_e) = \mathbf{H}^T\,\mathbf{F}(\boldsymbol{\theta})\mathbf{H}$, which can be expended as

$$\mathbf{J}_{geo}(\mathbf{p}_e) = \sum_{m=1}^{M}\mathbf{G}_m^T\,\mathbf{F}_m\mathbf{G}_m + \sum_{m=1}^{M}\sum_{n=m+1}^{M}(\mathbf{G}_m+\mathbf{G}_n)^T\,\mathbf{CI}_{mn}(\mathbf{G}_m+\mathbf{G}_n) \tag{4-10}$$

The matrices are defined in (3-14) and (3-20). The first summation in (4-10) is the combination of each pair independently and the second one is due to the cross term FIM, if there is some sensors participated in more than one pair.

We need to optimize CRLB, which is the inverse of FIM. The expression of inverse of (4-10) is complicated even if there is only the first summation. So instead of discussing

the optimal eigenvalues of CRLB, we will discuss the eigenvalues of $\mathbf{J}_{geo}$ to avoid the complexity of calculating $\mathbf{J}_{geo}^{-1}$. Define the eigen-decomposition of $\mathbf{J}_{geo}$ as

$$\mathbf{J}_{geo} = \sum_{i=1}^{p} \frac{1}{\lambda_i} \mathbf{v}_i^T \mathbf{v}_i \tag{4-11}$$

To increase the accuracy of the estimates is equal to reduce the size of the error ellipse. But which measurement we should use to be our criterion? Since the CRLB is the inverse of FIM, based on matrix theory; the eigenvalues of them are reciprocal of each other. Then

$$\mathbf{CRLB}_{geo} = \mathbf{J}_{geo}^{-1} = \sum_{i=1}^{p} \lambda_i \mathbf{v}_i^T \mathbf{v}_i \tag{4-12}$$

Since $\mathbf{J}_{geo}$ is positive definite matrix, all its eigen-values $1/\lambda_i$ are positive. It is easy to show that minimizing $\det(\mathbf{J}_{geo}^{-1}) = \prod_{i=1}^{p} \lambda_i$ is equivalent to maximizing $\det(\mathbf{J}_{geo}) = \prod_{i=1}^{p} 1/\lambda_i$. So the decrease of the area of CRLB ellipse is equal to the increase of the FIM ellipse.

Is that also true for the perimeter? Let $\text{trace}(\mathbf{J}_{geo}) = \sum_{i=1}^{p} 1/\lambda_i$ and $\text{trace}(\mathbf{J}_{geo}^{-1}) = \sum_{i=1}^{p} \lambda_i$. Using the Cauchy-Schwarz inequality [3] gives

$$p = \begin{bmatrix} \sqrt{\lambda_1} & \cdots & \sqrt{\lambda_p} \end{bmatrix} \begin{bmatrix} \sqrt{\lambda_1} \\ \vdots \\ \sqrt{\lambda_p} \end{bmatrix} \le \sqrt{\sum_{i=1}^{p} \lambda_i} \sqrt{\sum_{i=1}^{p} 1/\lambda_i} = \sqrt{\text{tr}(\mathbf{J}_{geo})\text{tr}(\mathbf{J}_{geo}^{-1})} \tag{4-13}$$

This gives $\text{tr}(\mathbf{J}_{geo}^{-1}) \ge p^2 / \text{tr}(\mathbf{J}_{geo})$. Thus the maximization of $\text{tr}(\mathbf{J}_{geo})$ tends to bring the minimization of $\text{tr}(\mathbf{J}_{geo}^{-1})$.

Based on matrix theory and (4-10)

$$\mathrm{tr}(\mathbf{J}_{geo}) = \sum_{m=1}^{M} \mathrm{tr}(\mathbf{J}_m) + \sum_{m=1}^{M} \sum_{n=m+1}^{M} \mathrm{tr}(\mathbf{J}_{m,n}) \tag{4-14}$$

where $\mathbf{J}_m = \mathbf{G}_m^T \mathbf{F} \mathbf{I}_m \mathbf{G}_m$ and $\mathbf{J}_{m,n} = (\mathbf{G}_m + \mathbf{G}_n)^T \mathbf{C} \mathbf{I}_{mn} (\mathbf{G}_m + \mathbf{G}_n)$. So if we choose trace of $\mathbf{J}_{geo}$ as our optimal criterion, we can discuss pair-wise.

But the determinant of $\mathbf{J}_{geo}$ generally does not have the formula as (4-14), there is not an explicit expression of the determinant of the summation of matrices. The Minkowski Inequality [3] says for (4-10)

$$\det(\mathbf{J}_{geo}) \geq \sum_{m=1}^{M} \det(\mathbf{J}_m) + \sum_{m=1}^{M} \sum_{n=m+1}^{M} \det(\mathbf{J}_{m,n}) \tag{4-15}$$

with equality if and only if all the $\mathbf{J}_m$ and $\mathbf{J}_{m,n}$ are positively linearly dependent. It is complicated if we can not separate pairs.

From the above discuss, there is a trade off between choosing the determinant or the trace of $\mathbf{J}_{geo}$ to be our optimal criterion. (i) For determinant, maximize the determinant of $\mathbf{J}_{geo}$ definitely minimize the determinant of $\mathbf{J}_{geo}^{-1}$, but the expression of $\det(\mathbf{J}_{geo})$ with respect to the parameters which we are going to optimal is complicated and hard to discuss; (ii) For trace, based on (4-14), the trace of $\mathbf{J}_{geo}$ can be separated pair-wised, which makes the optimal method done inside each pair, but the maximal of $\mathrm{tr}(\mathbf{J}_{geo})$ can not guarantee the minimal of $\mathrm{tr}(\mathbf{J}_{geo}^{-1})$, even though it always has the tendency. We choose trace as our criterion, since normally the structure of $\mathbf{J}_{geo}$ will satisfy the maximization of $\mathrm{tr}(\mathbf{J}_{geo})$ bring the minimization of $\mathrm{tr}(\mathbf{J}_{geo}^{-1})$, and pair-wised optimization brings more flexibility.

4.2 Diversity of Sensor Selection and Pairing

Each element of Fisher Information Matrix $\mathbf{J}_{geo}$ is a function of both emitter location $\mathbf{p}_e$ and sensor navigation data $\mathbf{s}$ and $\dot{\mathbf{s}}$, so is the trace of $\mathbf{J}_{geo}$. Our objective is to select an optimal subset of sensors and pair them as well. The optimal function is given by

$$\{\text{subset of sensors}\}^* = \arg\max_{\text{subset}} \left\{\text{trace}\left(\mathbf{J}_{geo}\left(\text{subset}\right)\right)\right\} \qquad (4\text{-}16)$$

The computation of $\mathbf{J}_{geo}$ depends on how the sensors are cooperated. There are many ways to manage a set of sensors according to the sensor network types. We propose various approaches to this problem and discuss trade-offs between them, for example (i) assumes that the sensors have pre-paired and share their data between these pairs; sensor selection then consists of selecting pairs to optimize performance while meeting constraints on number of pairs selected; (ii) consists of optimally determining pairings as well as selections of pairs but with the constraint that no sensors are shared between pairs; (iii) consists of allowing sensors to be shared between pairs; and etc.

4.2.1 Network Types

Wireless sensor network is the network that sensors are cooperatively monitor physical or environmental conditions. So the sensors need to be grouped together under some constraints. The TDOA/FDOA measurements need a set of sensors selected and paired. There are many ways to cooperate the sensors and let them paired. The sensors are connected according to some requirements, such as distance, effective signal noise ratio, efficiency and etc. We define the three normally practical types of sensor network as following and illustration is in Figure 10.

Type-I: No Sensor Sharing (a pair that does not share any of its sensors with any other pairs is called "*independent pair*".);

Type-II: De-Centralized Sensor Sharing (some sensors are participated in more than one pair, and there is no center sensor.);

Type-III: Centralized Sensor Sharing (a common reference/fusion/center sensor is used, all the other sensors are paired with it.)

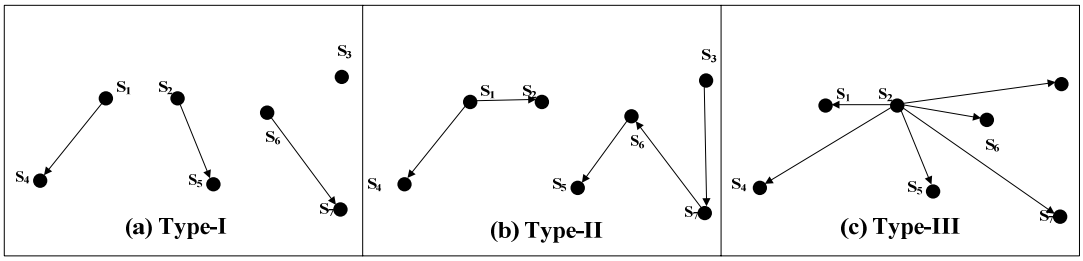

Figure 10 Three type of sensor networks

Based on the different network types, the FIM of geo-location will have different structure which is discussed later.

4.2.2 Two Pairing Scenarios

Besides the specified sensor network types we need to consider, there are also two different scenarios for the given sensors: (i) pre-paired sensors, in which all the sensors are already paired, we can not change their pairing topology; (ii) free sensors, a sensor that could be paired to any other sensors is called "*free sensor*", we can pair them with more flexibility. In the following sessions, we discussed the sensor selections algorithms for the three network types and two scenarios.

4.3 Sensor Selection and Pairing Strategies

In the previous session, we have discussed about the diversity of pairing strategies, in this session we give solutions for them.

4.3.1 Pre-Paired Sensors

When sensors are pre-paired, since we could not depart the pairs and re-pair them, we select pairs instead of sensors. We do not have much flexibility to pair the sensors. The TDOA/FDOA pair-FIM $\mathbf{F}_m$ and cross-FIM $\mathbf{CI}_{m,n}$ are evaluated based on the paring and sensor sharing topology.

(1) Type-I: No Sensor Sharing

When no sensor is shared all the cross-FIM $\mathbf{CI}_{m,n}$ is zero. The $\mathbf{F}_m$ is evaluated individually for each pair. Then $\mathbf{F}_\mathbf{0}$ of all TDOA/FDOA estimation will have the block diagonal structure as

$$\mathbf{F}_\mathbf{0} = \begin{bmatrix} \mathbf{F}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{F}_2 & \ddots & \mathbf{0} \\ \vdots & \ddots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{F}_M \end{bmatrix} \tag{4-17}$$

FIM of geo-location (4-10) becomes

$$\mathbf{J}_{geo}(\mathbf{p}_e) = \sum_{m=1}^{M} \mathbf{G}_m^T \mathbf{F}_m \mathbf{G}_m \tag{4-18}$$

The problem of selecting *K* pairs from *M* pairs is specified by

$$\begin{aligned} \{p_1, \cdots, p_M\}^* &= \arg \min_{p_1, \cdots, p_M} \left\{ \text{trace}\left[p_1 \cdot \left(\mathbf{G}_1^T \mathbf{F}_1 \mathbf{G}_1 \right) + \cdots + p_M \cdot \left(\mathbf{G}_M^T \mathbf{F}_M \mathbf{G}_M \right) \right] \right\} \\ s.t. \quad & p_1 + \cdots p_m \cdots + p_M = K < M, \quad p_m \in \{0,1\} \end{aligned} \tag{4-19}$$

It is to select the *K* pre-paired sensor pairs that have the largest values of $\text{trace}(\mathbf{J}_m)$.

To demonstrate the capability of the FIM-based pair selection methods we present the simulation result for the case of locating an emitter with a random lay-down of 10 pair of sensors as in Figure 11. The system arrangement is: the emitter is located at the origin, sensors are randomly placed in the range of $20km \times 10km$, the distance of each pair is under $2km$, and the norm of velocity is smaller than $300km/s$.

The sensor selection proceeds as follows. Each sensor intercepts the emitter signal data at signal to noise ratio in the range of 10~15dB (where the SNR variation is assumed to depend quadratic ally on the range to the emitter). The full set of sensors share a very small amount of data to obtain a rough estimate of the emitter location; alternatively, we could consider the case where the system is cued by some other sensor system that provides a rough location that is to be improved using our sensors. Based on the initial location estimation and small amount of data, a central sensor (or other control center) will do the selection job. The selected pairs then participate the following tasks.

Figure 12 shows the performance of sensor selections without sensor sharing for the pre-paired case. We select 2 to 10 pairs shown on the horizontal axis. The vertical axis shows the normalized dB value standard deviation (normalized by the largest standard deviation) of the estimated geo-location distance error versus the number pairs selected. It is the trace of the covariance matrix of estimation errors, which is the variance of the distance from the estimation to the true emitter location physically. The upper curve (-Δ-) shows the performance for the original selected pairs; the lower curve (-O-) shows the performance when using the FIM-based selection method discussed above. Not surprisingly, the ability to select the pairing on the basis of the sensor geometry and the

rough emitter location enables better performance. And for the layout as in our simulation, only 2 pairs selection can achieve the accuracy as the 8 pairs selected originally.

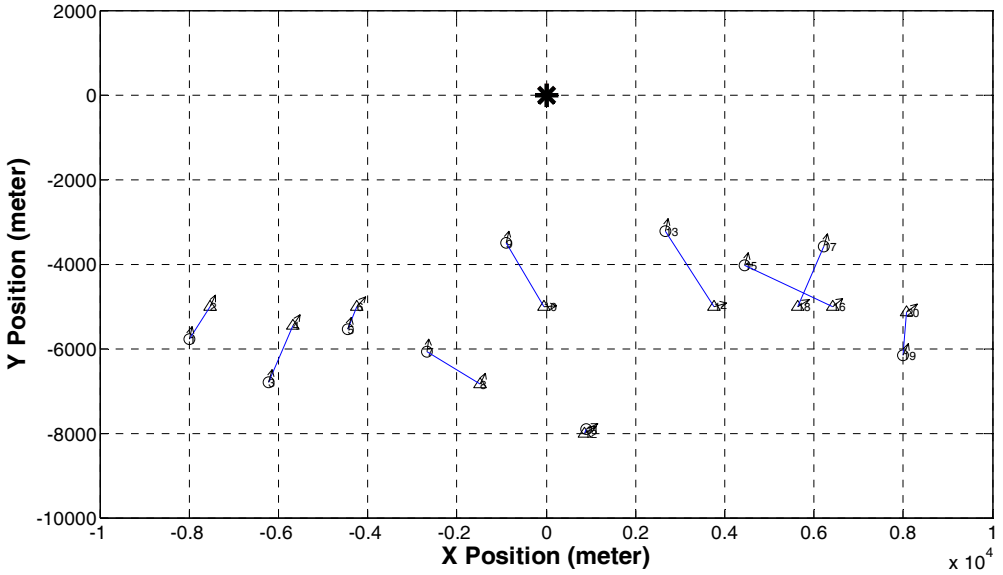

Figure 11 Layout of 10 pairs of sensor

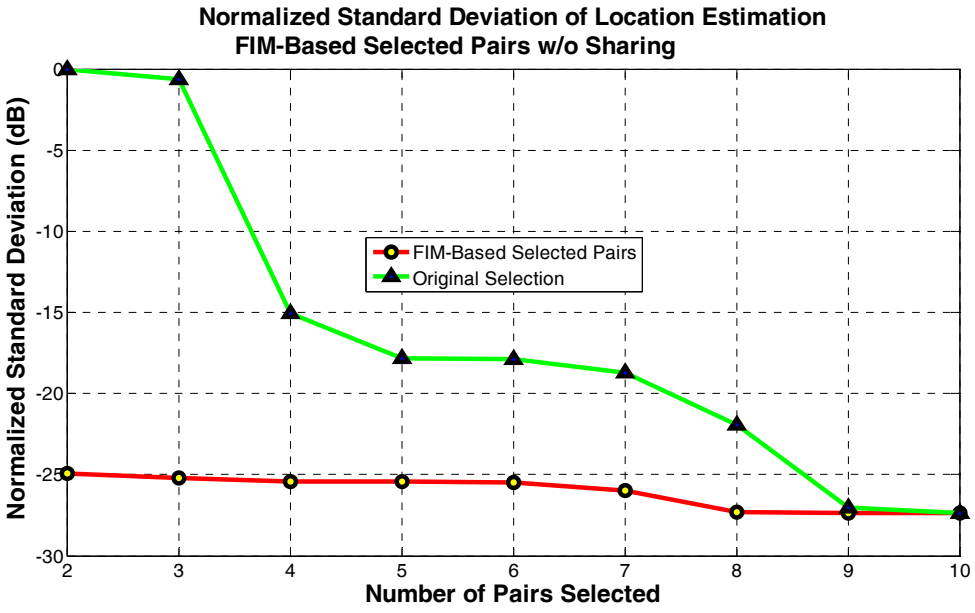

Figure 12 Simulation result for FIM-based sensor pairs selection w/o sharing. The Standard deviation is normalized and taken the dB value.

(2) Type-II: De-Centralized Sensor Sharing

Under this type, the network looks messy. Some pairs are independent; some sensors are participated in more than one pair. For type-I, we can treat the sensors pair by pair, since there is no correlation when estimating TDOA/FDOAs. Here we define a group of sensors as a sensor set by: A set is defined by a group of sensors which do not have connections with any other sensors outside the group and do not have independent pairs inside the group. For the sensor network in Figure 10-(b), the sets are defined as in Figure 13. Set-1 is grouped by sensor-1,2, and 4; Set-2 is grouped by sensor-3,5,6 and 7.

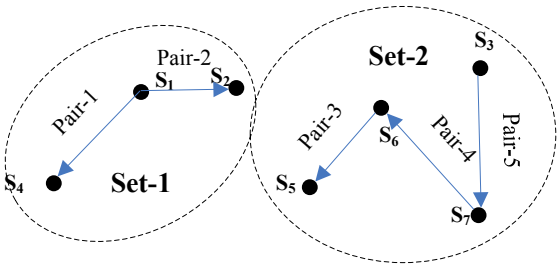

Figure 13 Sensor sets example

We calculate the FIM of geo-location of each set. For example, the evaluation of set-1 is

$$\mathbf{J}_{geo,set-1} = \mathbf{G}_1^{\mathrm{T}}\mathbf{F}_1\mathbf{G}_1 + \mathbf{G}_2^{\mathrm{T}}\mathbf{F}_2\mathbf{G}_2 + (\mathbf{G}_1+\mathbf{G}_2)^T\mathbf{C}\mathbf{I}_{1,2}(\mathbf{G}_1+\mathbf{G}_2) \qquad (4\text{-}20)$$

The problem of selecting K sensors from M sets is specified by

$$\begin{aligned} &\max_{p_1,\cdots,p_N}\{trace(p_1\cdot\mathbf{J}_{geo,set-1}+\cdots+p_M\cdot\mathbf{J}_{geo,set-M})\} \\ &s.t.\quad p_1\cdot n_1+\cdots+p_1\cdot n_M = K < N,\quad p_i\in\{0,1\} \\ &\qquad n_i \text{ is the number of sensors in set-i} \end{aligned} \qquad (4\text{-}21)$$

The simulation procedure and result are similar to the type-I, so we skip the result figures.

(3) Type-III: Centralized Sensor Sharing

For the pre-paired case, the central sensor is already specified and the remaining *N*-1 sensors pair with it to form *N*-1 centralized pairs. The center one must be selected. How to select the remaining *K*-1 sensors? They are all correlated when estimating TDOA/FDOAs for each pair. There are C_{N-1}^{K} possible ways to select *K* pairs. For each selection, for illustration renumber the pair number as in Figure 14.

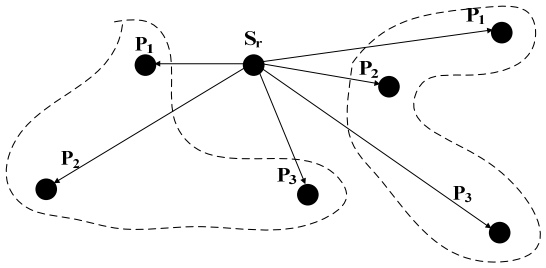

Figure 14 For each group, renumber each pair

The FIM of each group will have the following structure

$$\mathbf{J}_{geo,k} = \sum_{k=1}^{K} \mathbf{G}_k^{\mathrm{T}} \mathbf{F}_k \mathbf{G}_k + \sum_{m=1}^{K} \sum_{n=m+1}^{K} (\mathbf{G}_m + \mathbf{G}_n)^T \mathbf{C}\mathbf{I}_{mn} (\mathbf{G}_m + \mathbf{G}_n) \qquad (4\text{-}22)$$

Assume r^{th} sensor is the reference one. Also the simulation procedure and result are similar to the type-I, so we skip the result figures.

From all above, for the pre-paired sensor networks, all we can do is calculating the FIM following some rules and select pairs. There is not much flexibility when doing the sensor selection and pairing.

4.3.2 Free Sensors

We are given a set of sensors and asked to optimally choose a subset and the optimal pairing them as well. In this case the pairing provides more flexibility to enable better performance but it introduces additional complexity as well. We develop the solutions for both no sensor sharing case and allowed sensor sharing case.

(1) No Sensor Sharing

Pair the free sensors without sharing any one of them. For N sensors, there could be $N/2$ independent pairs. To choose K ($\leq N/2$) pairs is a time-consuming work if we enumerated all the possible solutions. For example, N=10, there are $C_{10}^{2} = 45$ possible pairs, and $(N-1)\cdot(N-3)\cdots 3\cdot 1 = 945$ possible ways to make 5 pairs as a subset. Another example, for sensor-1,2,3 and 4, the 2 pairs could have the following solutions:

(i) (1,2) and (3,4)

(ii) (1,3) and (2,4)

(iii) (1,4) and (2,3)

Fortunately, since there is no sensor sharing and we select sensors pair by pair, the selection of next pair will not affect the selection of the previous selected pairs. It is like a kind of tree structure. We can use the Integer Dynamic Programming method to optimal the selection. For this paper, we used 'Branch and Bound' [5] method to choose one pair at each step.

The objective function is as

$$\max_{\text{all feasible solutions}} \left\{ \sum_{k=1}^{K} \text{trace}(\mathbf{J}_{\text{geo, k}^{\text{th}}\text{pair in the solution}}) \right\} \qquad (4\text{-}23)$$

Feasible solution means no sensor sharing in the solution and the number of sensors is as required.

An detailed example of branch and bound method is given in more detail in Appendix-4A.

Simulation results for the case of locating an emitter with a random lay-down of 14 sensors, all the parameters used are as in the pre-paired case. Figure 12 shows the

performance of sensor selection and pairing without sensor sharing for the free sensors case. We select 2 to 7 pairs, shown on the horizontal axis. The vertical axis shows the normalized dB value standard deviation of the geo-location distance error standard derivation versus the number pairs selected. The upper curve (-Δ-) shows the performance for the selected pairs under natural order as (1,2), (3,4) and (5,6)……; the lower curve (-O-) shows the performance when using the FIM-based pairing method discussed above.

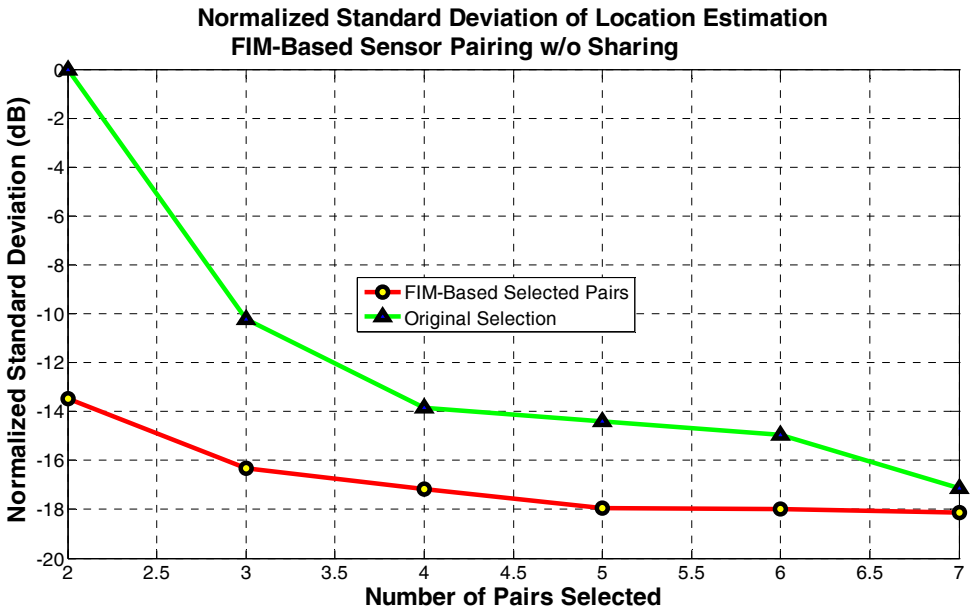

Figure 15 Simulation result for 14 free sensors selection and pairing.

The standard deviation is normalized and taken the dB value.

(2) Allowed Sensor Sharing

In the sensor sharing case, some sensors will work in more than one pairs to make more pairs and more TDOA/FDOA measurements. For N sensors, there could be C_N^2 possible pairs. To choose K $(\leq C_N^2)$ pairs, there could be $C_{C_N^2}^K$ possible ways. For example,

for N=10 and K=5, the solution number is 122,1759. That is a huge number, definitely impossible to list all of them. But fortunately, among all the solutions only a small part of them are useful or unique. We have established the following theorem and proved it.

Theorem: For M sensors, at most M-1 pairs can be used as a group; and different pairing method of the M sensors to make M-1 pairs will result in the same Cramer-Rao lower bound of geo-location estimation.

Prove:

(1) For M sensors, at most M-1 pairs are linearly independent.

Assume N sensors construct M pairs. Let r_k be the distance between sensors S_k and emitter, and $d_m = r_{m_1} - r_{m_2}$ be the range difference of the m^{th} $(1 \le m \le M)$ pair, the subscript m_1 and m_2 mean that the pair is by $(m_1)^{th}$ and $(m_2)^{th}$ sensors, for $m_1, m_2 \in \{1, 2, ...N\}$, and $m_1 \ne m_2$.

Let $\mathbf{rd} = [d_1 \quad \cdots \quad d_m \quad \cdots \quad d_M]^T$ and $\mathbf{r} = [r_1 \quad \cdots \quad r_n \quad \cdots \quad r_N]^T$. Then the relationship between $\mathbf{rd}$ and $\mathbf{r}$ is given by

$$\mathbf{rd}_{M\times 1} = \mathbf{T}_{M\times N} \cdot \mathbf{r}_{N\times 1} \tag{4-24}$$

where $\mathbf{T}$ is a sparse $M \times N$ matrix, which has only one '1' and one '-1' at each row. For the structure of $\mathbf{T}$, the rank of $\mathbf{T}$ is less than N-1, which means there are at most N-1 d_m linearly independent. For example, the pairing in Figure 16: (i) pairs (1,2), (2,4) and (1,4) are linearly dependent pairs, for these three sensors (4-24) is given as

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} r_1 \\ r_2 \\ r_4 \end{bmatrix} \tag{4-25}$$

(ii) pairs (3,5), (5,6) and (6,7) are linearly independent pairs, for these pairs (4-24) is given as

$$\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} r_3 \\ r_5 \\ r_6 \\ r_7 \end{bmatrix} \tag{4-26}$$

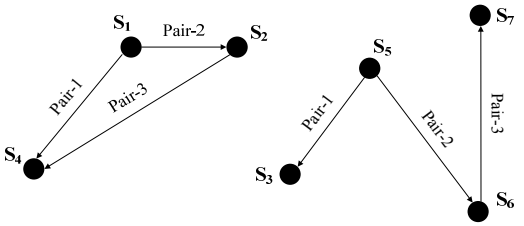

Figure 16 An example of linearly independent and dependent pairs

Linearly independent group, simply speaking, there are no closed loops for any sensors in the group. For *N* sensors, if we connect them one by one following the nature order, then for a open loop we can get *N*-1 pairs and for a close loop N pairs. At first glance, it seems like one more pair of TDOA/FDOA measurements will increase the accuracy of the estimation of emitter location. Why can not we use the close loop? Assume the $\mathbf{rd} = [d_1 \quad \cdots \quad d_M]^T$ is an open loop, and $\mathbf{rd}' = \left[\mathbf{rd}^T \quad d_{M+1}\right]^T$ is a close loop, then there is a relationship exist as

$$d_{M+1} = \boldsymbol{\beta}^T \cdot \mathbf{rd} \tag{4-27}$$

where $\boldsymbol{\beta}$ is a $M \times 1$ vector with only '1', '-1' or '0' as its elements.

Since the TDOA estimates are obtained by Maximum Likelihood estimator and $d_m = c\tau_m$, then

$$\Delta d_{M+1} = \boldsymbol{\beta}^T \cdot \Delta \mathbf{rd} \tag{4-28}$$

where Δd_{M+1} and $\Delta\mathbf{d}$ are the estimation errors. The signal models used in the second stage estimation will be like

$$\begin{aligned}\mathbf{d} &= \mathbf{f}(\mathbf{p}_e) + \Delta\mathbf{d} \\ d_{M+1} &= f_{M+1}(\mathbf{p}_e) + \boldsymbol{\beta}^T \cdot \Delta\mathbf{d}\end{aligned} \tag{4-29}$$

Therefore the error vector $\mathbf{e}' = [\Delta\mathbf{d}^T \quad \boldsymbol{\beta}^T \cdot \Delta\mathbf{d}]^T$ is not i.i.d (independent identical distribution), we can not use the Least Square method to estimate $\mathbf{p}_e$. Therefore the one more pair is not able to be used. Or on the other hand, the usage of the one more TDOA/FDOA measurements will not give any more benefit at the second stage estimation, all its information is embedded in other pairs.

(2) Unique CRLB for a group

For *M* sensors, we can make as many as *M*-1 pairs without a close loop. For network type-III in Figure 10, there is *M*-1 ways to make these *M*-1 pairs. We are going to prove that the CRLB is independent of the choice of reference sensor.

Let there be a different reference sensor (RS) in each different linearly independent group as in Figure 17. But the sensors inside each group are the same. In group-I, S_3 is the RS; in group-J, S_2 is the RS.

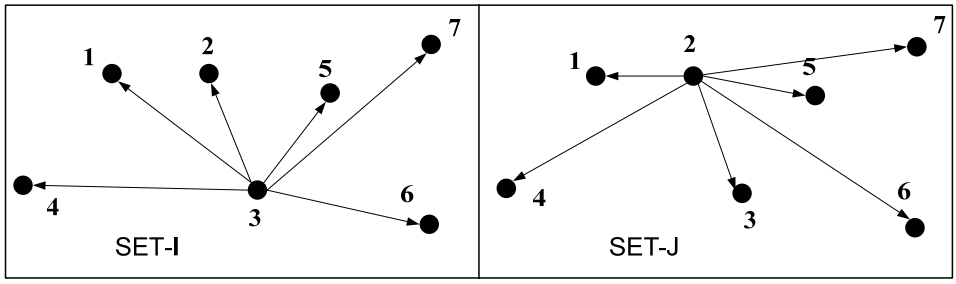

Figure 17 An example of different choice of reference sensor

Let c be the signal propagation speed, the range difference equations is

$$r_{kj} = c\tau_{kj} = r_k - r_j \tag{4-30}$$

When the receivers are moving, taking time derivative of (4-30) yields a set of FDOA measurement equation

$$\dot{r}_{kj} = c\dot{\tau}_{kj} = \dot{r}_k - \dot{r}_j \tag{4-31}$$

where $\dot{r}_i$ is the rate of change of r_i. From the time derivative of (3-15), $\dot{r}_i$ is related to the unknown location $\mathbf{p}_e$ by

$$\dot{r}_i = \frac{(\mathbf{s}_i - \mathbf{p}_e)^T \dot{\mathbf{s}}_i}{r_i} \tag{4-32}$$

Let $\mathbf{r} = [r_1, r_2, \cdots, r_N, \dot{r}_1, \dot{r}_2, \cdots, \dot{r}_N]^T$, $\mathbf{p}_i = [r_{1i}, \cdots, r_{ki}, \cdots, r_{Ni}; \dot{r}_{1i}, \cdots, \dot{r}_{ki}, \cdots, \dot{r}_{Ni}]^T, k \neq i$ and $\mathbf{p}_j = [r_{1j}, \cdots, r_{kj}, \cdots, r_{Nj}, \dot{r}_{1j}, \cdots, \dot{r}_{kj}, \cdots, \dot{r}_{Nj}]^T, k \neq j$ then

$$\mathbf{p}_i = \begin{bmatrix} \mathbf{T}_i & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_i \end{bmatrix} \cdot \mathbf{r} = \mathbf{H}_i \cdot \mathbf{r} \tag{4-33}$$

and

$$\mathbf{p}_j = \begin{bmatrix} \mathbf{T}_j & \mathbf{0} \\ \mathbf{0} & \mathbf{T}_j \end{bmatrix} \cdot \mathbf{r} = \mathbf{H}_j \cdot \mathbf{r} \tag{4-34}$$

where $\mathbf{T}_i$ is a $(N-1) \times N$ matrix, and has the following structure

$$\left[\mathbf{T}_i\right]_{m,n} = \begin{cases} -1, \ n = i \\ 1, \ m = 1, 2, ..., i-1, \ n = m \\ 1, \ m = 1, 2, ..., i-1, \ n = m+1 \\ 0, \ \text{others} \end{cases} \tag{4-35}$$

$\mathbf{0}$ is a $(N-1) \times N$ matrixes with all 0 entries. Let a $2(N-1) \times 2(N-1)$ full rank matrix $\mathbf{Q}_{ij}$ has the structure as

$$\mathbf{Q}_{ij} = \begin{bmatrix} \mathbf{E}_{ij} & \mathbf{0} \\ \mathbf{0} & \mathbf{E}_{ij} \end{bmatrix} \tag{4-36}$$

where $\mathbf{E}_{ij}$ is the elementary matrix, which is generated by exchanging the i^{th} and j^{th} rows of the identity matrix. So $\mathbf{Q}_{ij}$ will satisfy

$$\mathbf{H}_j = \mathbf{Q}_{ij} \cdot \mathbf{H}_i \tag{4-37}$$

then

$$\mathbf{p}_j = \mathbf{H}_j \cdot \mathbf{r} = \mathbf{Q}_{ij} \cdot \mathbf{H}_i \cdot \mathbf{r} = \mathbf{Q}_{ij} \cdot \mathbf{p}_i = \mathbf{g}(\mathbf{p}_i) \tag{4-38}$$

Based on the vector parameter CRLB transformation property [1], $\text{CRLB}(\mathbf{p}_j)$ and $\text{CRLB}(\mathbf{p}_i)$ has the following relationship

$$\text{CRLB}(\mathbf{p}_j) = \left[\frac{\partial \mathbf{g}(\mathbf{p}_i)}{\partial \mathbf{p}_i}\right] \text{CRLB}(\mathbf{p}_i) \left[\frac{\partial \mathbf{g}(\mathbf{p}_i)}{\partial \mathbf{p}_i}\right]^T = \mathbf{Q}_{ij} \text{CRLB}(\mathbf{p}_i) \mathbf{Q}_{ij}^T \tag{4-39}$$

Denote the emitter location $\mathbf{p}_e$ estimated by sensor pair group-I as $\hat{\mathbf{p}}_e^i$, the FIM of $\mathbf{p}_e^i$ is

$$\text{FIM}(\mathbf{p}_e^i) = \mathbf{H}_i^T \cdot \mathbf{C}_i^{-1} \cdot \mathbf{H}_i \tag{4-40}$$

where $\mathbf{C}_i$ is the covariance matrix of TDOA/FDOA measurements calculated by $\mathbf{C}_i = \text{E}[\hat{\mathbf{p}}_i \cdot \hat{\mathbf{p}}_i^T]$.

Since TDOA/FDOA are estimated by ML method, we assume the covariance of the estimates can reach CRLB, so $\mathbf{C}_i = \text{CRLB}(\mathbf{p}_i)$ and

$$\mathbf{C}_j = \text{CRLB}(\mathbf{p}_j) = \mathbf{Q}_{ij} \cdot \text{CRLB}(\mathbf{p}_i) \cdot \mathbf{Q}_{ij}^T = \mathbf{Q}_{ij} \cdot \mathbf{C}_i \cdot \mathbf{Q}_{ij}^T \tag{4-41}$$

$\mathbf{H}_j$ is the Jacobin matrix of group-J defined by

$$\mathbf{H}_j = \frac{\partial \mathbf{p}_j(\mathbf{p}_e)}{\partial \mathbf{p}_e} = \frac{\partial [\mathbf{Q}_{ij} \cdot \mathbf{p}_i(\mathbf{p}_e)]}{\partial \mathbf{p}_e} = \mathbf{Q}_{ij} \cdot \frac{\partial \mathbf{p}_i(\mathbf{p}_e)}{\partial \mathbf{p}_e} = \mathbf{Q}_{ij} \cdot \mathbf{H}_i \tag{4-42}$$

Then

$$
\begin{aligned}
\mathrm{FIM}(\mathbf{p}_e^j) &= \mathbf{H}_j^T \cdot \mathbf{C}_j^{-1} \cdot \mathbf{H}_j \\
&= (\mathbf{Q}_{ij} \cdot \mathbf{H}_i)^T \cdot (\mathbf{Q}_{ij} \cdot \mathbf{C}_i \cdot \mathbf{Q}_{ij}^T)^{-1} \cdot (\mathbf{Q}_{ij} \cdot \mathbf{H}_i) \\
&= \mathbf{H}_i^T \cdot [\mathbf{Q}_{ij}^T (\mathbf{Q}_{ij}^T)^{-1}] \cdot \mathbf{C}_i^{-1} \cdot [\mathbf{Q}_{ij}^{-1} \mathbf{Q}_{ij}] \cdot \mathbf{H}_i \\
&= \mathbf{H}_i^T \cdot \mathbf{C}_i^{-1} \cdot \mathbf{H}_i \\
&= \mathrm{FIM}(\mathbf{p}_e^i)
\end{aligned}
\qquad (4\text{-}43)
$$

We just proved that for a group of sensors, the FIM is identical (and so is the CRLB) if we choose network type-III to pair them, which selects one sensor as reference one. For the sensor network type-II, there is no reference sensor. We pair all of them and avoid any close loop. Then the matrix $\mathbf{T}_i$, defined as in (4-33), is a $(N-1)\times N$ matrix, which only has one '1' and one '-1' in each row, and so does $\mathbf{T}_j$, which is the transform matrix for another pairing method. The full rank matrix transform $\mathbf{Q}_{ij}$ does not have a obvious structure as in (4-35), but it is not hard to find the transform matrix and satisfies (4-37). The following proven is the same as we prove the type-III.

We get the final conclusion, for *M* sensors, we can have at most *M*-1 pairs usable, and the FIM/CRLB does not depend on the pairing method if there is not any close loop. □

Based on this statement, it is possible to make the sensor selection and pairing. When we are given N sensors and asked to make K pairs, there are many solutions for this network. We can use at least $K+1$ sensors to make it or at most $2K$ sensors to make all independent pairs. Since the main advantage to share sensors is to save some sensors energies, we should use the number of sensors as less as possible. So here we only choose $K+1$ sensors to make K pairs.

For example, for given $N=7$ and $K=3$ pairs needed, compute the CRLB of geo-location of all $C_7^4 = 35$ solutions, and find the one with the smallest trace. Inside each

solution, sensors are paired by sequence or paired by one reference sensor. For example, as in Figure 18, the solution set is $\{S_1, S_2, S_3, S_4\}$, the pairs are $(S_1, S_2), (S_2, S_3), (S_3, S_4)$.

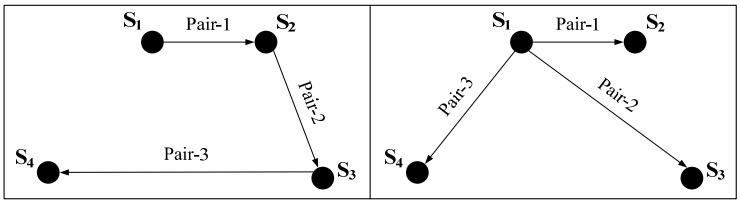

Figure 18 Example of sequence pairing and reference pairing

A simulation result for the case of locating an emitter with a random lay-down of 10 sensors (all the parameters used are as in the pre-paired case) is given in Figure 19. It shows the performance of sensor selection and pairing with sensor sharing for the free sensors case. For the given 10 sensors, we make 2 to 9 pairs, shown on the horizontal axis. The vertical axis shows the normalized dB value standard deviation of the geo-location distance error standard derivation versus the number pairs selected. The upper curve (-Δ-) shows the performance for some randomly selected pairs; the lower curve (-O-) shows the performance when using the FIM-based pairing method discussed above.

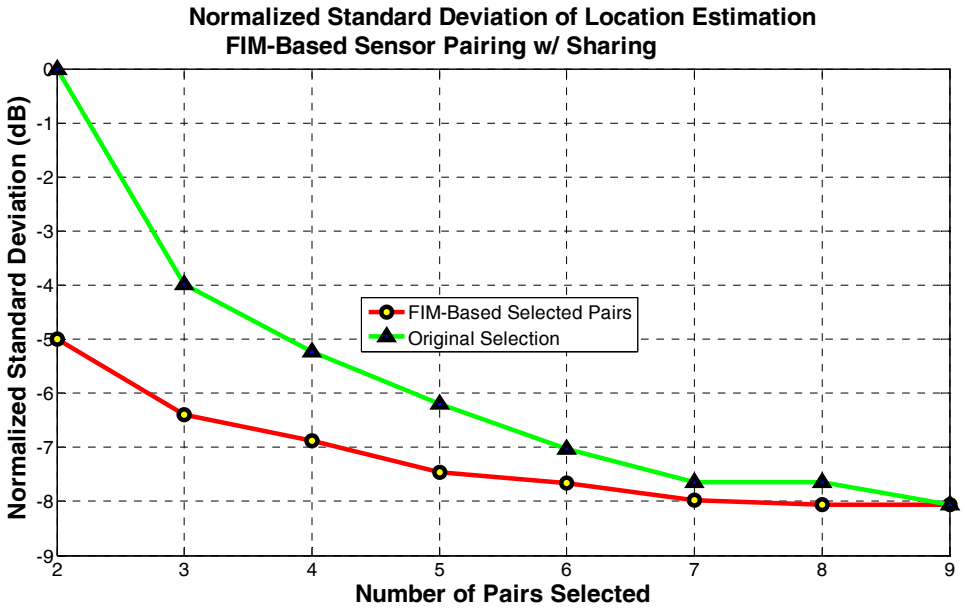

Figure 19 Simulation result 10 free sensors selection and pairing with sensor sharing. The standard deviation is normalized and taken the dB value.

4.3.3 Trade Off Between Non-Sharing and Sharing Methods

(1) Without Sensor Sharing

- FIM of Geo-Location is easy to calculate, since each pair is independent;
- However, the pairing method is more complicated, since we need to consider all the possible paring ways;
- From a system point of view, the communication among different pairs can be done simultaneously;
- The number of pairs needed is small; beyond a certain point the accuracy improves slowly as more pairs are selected to participate.

(2) With Sensor Sharing

- For a total of N sensors we can have as many as $N-1$ pairs, the more the higher accuracy of location estimation;
- Fortunately, FIM of all the possible independent sets are the same, so we do not need to consider about the pairing method. One simple way is to pair the sensors in nature order. This is the main result of this work and leads to a major reduction in the optimization processing required.
- However, since not all the pairs are uncoupled, there are cross terms in the TDOA/FDOA FIM. This complicates the computation required to support the optimization processing.
- Some sensors work in more than one pair, the communication among them needs to be considered carefully to avoid collision. This will be the focus of future work.

Appendix-4A an Example of Branch and Bound Method Used in Sensor Pairing

Branch and Bound method [5] is based on that the enumeration of integer solutions has a tree structure. At each step, a decision will be made on one integer variable. The main idea in branch and bound is to avoid growing the whole tree as much as possible. When a node is selected for further growing, it is like branching. Not all nodes are going to be evaluated, some are pruned in the intermediate steps.

Instead of enumerating all possible solutions, in each step to choose the next pair, we grow the only 'better' ones. By 'better', we mean whose have the higher probability to be optimal.

For $N = 8$, there are 28 possible pairs, randomly give each possible pair a value. (n, m) means the pair is paired be sensor-n and sensor-m. In the first step to choose the first pair, without loss of generality, we choose the pairs which have sensor-1 as one part. For example, if we choose pair (1,2) as the first pair, then the bounding function value for it is 71, which let pair (3,5)-(3,6)-(7,8) as the following choosing pairs. Since sensor-1 and sensor-2 are actually paired, we did not let them to reuse in the bounding function calculation at this node or any descendent nodes. The very best objective function value that we might have at a leaf node descended from (1,2)-?-?-? is 71. Sine sensor-3 is shared between two pairs, this solution is not feasible, but maybe needed to be fathomed, it depends on the incumbent we maybe find later. Based on this decision strategy, we got the bud nodes in the first step of the tree as Figure 20

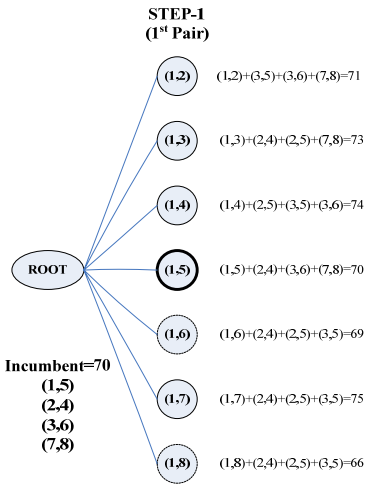

Figure 20 First step tree structure

The first step of the tree is generated from the root node by enumerating all the possible pairs which have sensor-1. By evaluate the bounding function, we get our first

best infeasible solution, (1,5)-(2,4)-(3,6)-(7,8) , as incumbent=70. And prune the pair(1,6) and (1,8), whose bounding function value are smaller than incumbent.

Pruned nodes are indicated by dashed border, incumbent node is by bold border, and infeasible nodes, whose bounding function value are larger than incumbent, are going to be growing in some order.

Since we are using the global-best node selection policy and the objective function is to maximize, we choose pair (1,7) , who has the largest bounding function value, for first further expansion, which gives the next tree step as Figure 21

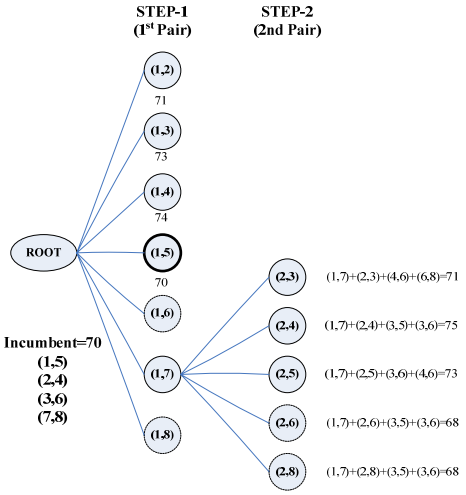

Figure 21 Second step tree structure

The expansion of pair(1,7) is at the second step of the tree. It is generated from the (1,7) node by enumerating all the possible pairs which have sensor-2, this is based on nature order. After evaluating all the bounding function values, some nodes were pruned. But we did not find a new incumbent in this expansion. Also the global-best node selection policy, after comparing all the bounding function values of current remaining nodes, partial solution (1,7)-(2,4) is to be fathomed next, which gives the result as Figure 22.

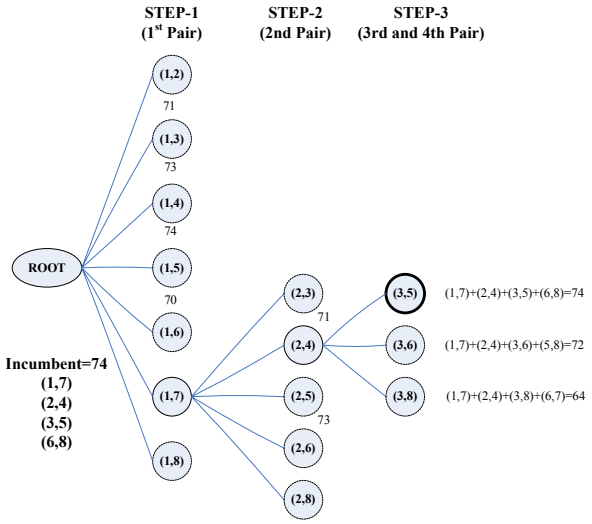

Figure 22 Third step tree structure

Since this is the last step in the tree, evaluating gives the objective function values. And we are lucky, since one value is larger than the incumbent, all the other leaf node values, and all the bounding values of infeasible bud nodes we got before. So we have the new incumbent, which is also the final solution as (1,7)-(2,4)-(3,5)-(6,8).

In this particular example, we also evaluated 15 nodes, which is much smaller than the work of a full enumeration of the 105 possible solutions.

Simulation in Figure 23 gives the comparison of time consuming between enumerating method and branch and bound method. We only let N=12 for the largest number, since for larger N, the enumerating way is even impossible to realize. The upper curve (-Δ-) shows the performance for the enumerating method; the lower curve (-O-) shows the performance when using the branch and bound method discussed above. The result shows that the dynamic programming method saves computation times largely.

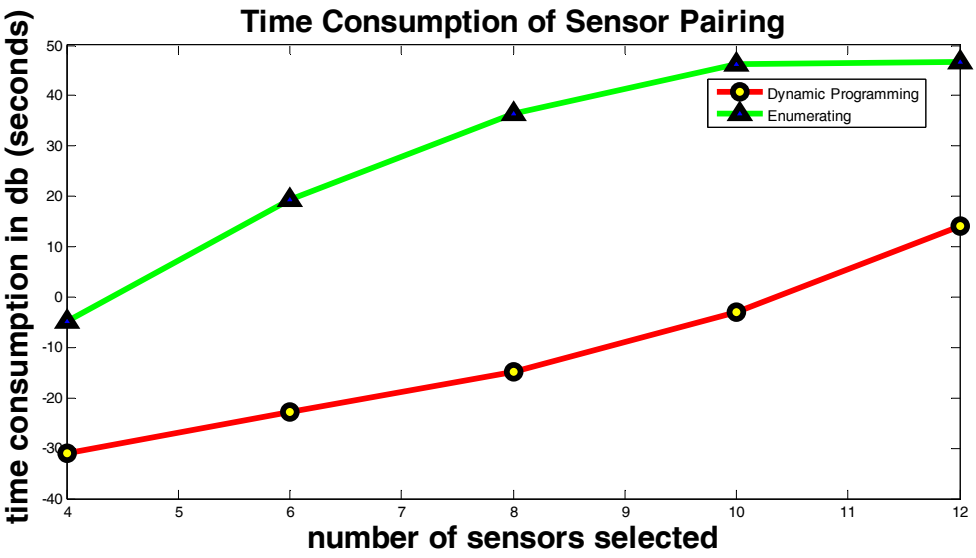

Figure 23 Simulation result for time consumption comparison between of 'branch and bound' method and enumerating method

5 Sensors' Navigation Data Error Effects and Mitigation

The signal model used in the second stage estimation depends on the sensors positions and velocity. And we have discussed in session 3.4 that the accuracy of emitter location is sensitive to the accuracy of sensors navigation data (NAV). Besides the TDOA/FDOA estimation errors, sensors' NAV errors can also lead to the emitter location errors. In practice, the sensor navigation data can not be known exactly, there is always error. This issue has been considered for a while, for example in [39], they analyzed the NAV errors effects and gave a closed form solution to solve this problem. But the algorithms are not simple and have many assumptions and simplifications. In this section, we will develop how the errors affect the estimation accuracy and the updated CRLB of geo-location. And we will take the navigation data errors into account when doing the second stage estimation to reach the updated CRLB.

5.1 Least-Squares Estimation Error Model

At the second stage estimation, we used non-linear least square method to estimate the geo-location. The signal model errors along with the measurement errors will impact in this stage.

Assume the m^{th} pair is paired by the (m,1)th and (m,2)th sensors, the signal model for TDOA τ_m and FDOA ω_m are given by

$$\begin{aligned}\tau_m &= \frac{1}{c}(\|\mathbf{p}_e - \mathbf{s}_{m,1}\| - \|\mathbf{p}_e - \mathbf{s}_{m,2}\|) \\ \upsilon_m &= \frac{f_e}{c}(\mathbf{u}_{m,1}^T \cdot \dot{\mathbf{s}}_{m,1} - \mathbf{u}_{m,2}^T \cdot \dot{\mathbf{s}}_{m,2})\end{aligned} \tag{5-1}$$

Define the sensors' parameter vector for the m^{th} pair as $\boldsymbol{\beta}_m = [\mathbf{s}_{m,1}^T, \dot{\mathbf{s}}_{m,1}^T, \mathbf{s}_{m,2}^T, \dot{\mathbf{s}}_{m,2}^T]^T$. Then (5-1) can be written as

$$\begin{aligned}\tau_m &= f_m^{\tau}(\mathbf{p}_e, \boldsymbol{\beta}_m) \\ \upsilon_m &= f_m^{\upsilon}(\mathbf{p}_e, \boldsymbol{\beta}_m)\end{aligned} \tag{5-2}$$

The TDOA/FDOA measurements are modeled as being corrupted by an additive noise vector $\mathbf{e}_m = [\Delta\tau_m, \Delta\upsilon_m]^T$. It also is not possible to know the sensors' position and velocity (the navigation data) without errors; thus they are modeled as being corrupted by an additive error vector $\boldsymbol{\eta}_m = [\Delta\mathbf{s}_{m,1}^T, \Delta\dot{\mathbf{s}}_{m,1}^T, \Delta\mathbf{s}_{m,2}^T, \Delta\dot{\mathbf{s}}_{m,2}^T]^T$. Vector forms of the quantities of interest at the total *M* pairs measurements are denoted as

$$\begin{aligned}\mathbf{B} &= \begin{bmatrix}\boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_M\end{bmatrix}, \quad \mathbf{f}(\mathbf{p}_e, \mathbf{B}) = \begin{bmatrix} f_1^{\tau}(\mathbf{p}_e, \boldsymbol{\beta}_1) \\ f_1^{\upsilon}(\mathbf{p}_e, \boldsymbol{\beta}_1) \\ \vdots \\ f_M^{\tau}(\mathbf{p}_e, \boldsymbol{\beta}_M) \\ f_M^{\upsilon}(\mathbf{p}_e, \boldsymbol{\beta}_M)\end{bmatrix} \\ \mathbf{e} &= \begin{bmatrix}\mathbf{e}_1 \\ \vdots \\ \mathbf{e}_M\end{bmatrix}, \quad \mathbf{N} = \begin{bmatrix}\boldsymbol{\eta}_1 \\ \vdots \\ \boldsymbol{\eta}_M\end{bmatrix}\end{aligned} \tag{5-3}$$

And then the TDOA/FDOA vector $\mathbf{m} = \begin{bmatrix}\tau_1 & \upsilon_1 & \cdots & \tau_M & \upsilon_M\end{bmatrix}^T$, measurement vector is $\hat{\mathbf{m}}$, and the sensors' navigation data estimation vector become, respectively as

$$\hat{\mathbf{m}} = \mathbf{f}(\mathbf{p}_e, \mathbf{B}) + \mathbf{e} \tag{5-4}$$

and

$$\hat{\mathbf{B}} = \mathbf{B} - \mathbf{N} \tag{5-5}$$

The emitter location problem then becomes: given TDOA/FDOA measurement vector $\hat{\mathbf{m}}$ and the local state estimate vector $\hat{\mathbf{B}}$, and also a reference location $\mathbf{p}_e^0$ (Which could be an estimate of $\mathbf{p}_e$ determined from a previous iteration of the estimation procedure or based upon a priori information.) compute an estimate $\hat{\mathbf{p}}_e$ of $\mathbf{p}_e$. The goal of this section is to characterize the impact of using $\hat{\mathbf{B}}$ in this processing rather than the unavailable actual $\mathbf{B}$.

Since $\mathbf{f}(\mathbf{p}_e, \mathbf{B})$ is a nonlinear vector function with respect to $\mathbf{p}_e$. To get a reasonably simple estimator, $\mathbf{f}(\mathbf{p}_e, \mathbf{B})$ can be linearized by expanding it in a Taylor expansion about $\mathbf{p}_e^0$ and retaining the first two terms as

$$\mathbf{f}(\mathbf{p}_e, \mathbf{B}) \cong \mathbf{f}(\mathbf{p}_e^0, \mathbf{B}) + \mathbf{H}(\mathbf{p}_e - \mathbf{p}_e^0) \tag{5-6}$$

where $\mathbf{H} = \left.\dfrac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{p}_e}\right|_{\mathbf{p}_e = \mathbf{p}_e^0}$. Combining (5-6) and (5-4) gives

$$\mathbf{r} = \mathbf{H} \cdot \mathbf{p}_e + \mathbf{e} \tag{5-7}$$

where $\mathbf{r} = \hat{\mathbf{m}} - \mathbf{f}(\mathbf{p}_e^0, \mathbf{B}) + \mathbf{H} \cdot \mathbf{p}_e^0$.

The least square approach is to solve

$$\hat{\mathbf{p}}_e = \arg\min_{\mathbf{p}_e} \left\{ \left[\mathbf{r} - \mathbf{H} \cdot \mathbf{p}_e \right]^T \mathbf{W} \left[\mathbf{r} - \mathbf{H} \cdot \mathbf{p}_e \right] \right\} \tag{5-8}$$

Thus the solution of (5-8) is

$$\begin{aligned} \hat{\mathbf{p}}_e &= \left(\mathbf{H}^T \mathbf{W} \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{W} \cdot \mathbf{r} \\ &= \mathbf{p}_e^0 + \left(\mathbf{H}^T \mathbf{W} \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{W} \left[\hat{\mathbf{m}} - \mathbf{f}(\mathbf{p}_e^0, \mathbf{B}) \right] \end{aligned} \tag{5-9}$$

Substituting (5-4) into (5-9) and rearranging terms, the expression of $\hat{\mathbf{p}}_e$ can be written in the form

$$\hat{\mathbf{p}}_e = \mathbf{p}_e + \left(\mathbf{H}^T \mathbf{W} \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{W} \left(\Delta \mathbf{f}_H + \mathbf{e} \right) \tag{5-10}$$

where $\Delta\mathbf{f}_H = \mathbf{f}(\mathbf{p}_e, \mathbf{B}) - \mathbf{f}(\mathbf{p}_e^0, \mathbf{B}) - \mathbf{H}(\mathbf{p}_e - \mathbf{p}_e^0)$ is the signal model linearization error.

We assume the exact knowledge of system parameter vector $\mathbf{B}$ at the above analysis up to (5-10). If $\hat{\mathbf{B}}$ is sufficiently close to $\mathbf{B}$, then the Taylor expansion with respect to $\mathbf{B}$ yields

$$\mathbf{f}(\mathbf{p}_e, \mathbf{B}) \cong \mathbf{f}(\mathbf{p}_e, \hat{\mathbf{B}}) + \mathbf{K}(\mathbf{B} - \hat{\mathbf{B}}) \tag{5-11}$$

where $\mathbf{K} = \left.\dfrac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{B}}\right|_{\mathbf{B}=\hat{\mathbf{B}}}$. Substitute (5-11) into (5-9) and rearranging terms, the expression of $\hat{\mathbf{p}}_e$ can be written in the form

$$\hat{\mathbf{p}}_e = \mathbf{p}_e + \left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\left(\Delta\mathbf{f}_H + \Delta\mathbf{f}_K + \mathbf{e}\right) \tag{5-12}$$

where $\Delta\mathbf{f}_K = \mathbf{K}(\mathbf{B} - \hat{\mathbf{B}}) = \mathbf{K}\mathbf{N}$ is the error caused by the sensors' navigation data (system parameters) uncertainty.

It is clear that the estimation error is composed by there error items: (i) linearization error $\mathbf{n}_1 = \Delta\mathbf{f}_H$; (ii) the signal models error $\mathbf{n}_2 = \Delta\mathbf{f}_K$; (iii) the TDOA/FDOA measurement error $\mathbf{n}_3 = \mathbf{e}$.

The bias of the estimator is

$$\mathrm{E}\left\{\hat{\mathbf{p}}_e - \mathbf{p}_e\right\} = \left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W} \cdot \mathrm{E}\left\{\Delta\mathbf{f}_H + \Delta\mathbf{f}_K + \mathbf{e}\right\} \tag{5-13}$$

And the covariance matrix of estimation error is given as

$$\mathrm{var}\left\{\hat{\mathbf{p}}_e - \mathbf{p}_e\right\} = \left[\left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\right] \cdot \mathrm{var}\left\{\Delta\mathbf{f}_H + \Delta\mathbf{f}_K + \mathbf{e}\right\} \cdot \left[\mathbf{W}\mathbf{H}\left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\right] \tag{5-14}$$

Since we use the iterative method to update the estimation, the linearization error $\Delta\mathbf{f}_H$ can be ignored at the final step. For the signal model error, there are two cases we need to discuss about: (i) If $\mathbf{B}$ is nonrandom, then the uncertainty of it will contribute to the bias of the estimator, since $\mathrm{E}\left\{\Delta\mathbf{f}_K\right\} = \Delta\mathbf{f}_K$. But it will not affect the covariance matrix of

estimation error, since $\text{var}\{\Delta\mathbf{f}_K\} = \mathbf{0}$; (ii) If $\mathbf{B}$ is random and has zeros mean and covariance matrix $\mathbf{C}_\mathbf{N}$, then the estimation bias only depend on $\mathbf{e}$. And assume $\mathbf{N}$ is independent with $\mathbf{e}$, let $\mathbf{W}^{-1} = \mathbf{C}_\mathbf{e} = \text{var}\{\mathbf{e}\}$, then the estimation error covariance matrix become to

$$\begin{aligned}\text{var}\{\hat{\mathbf{p}}_e - \mathbf{p}_e\} &= \left(\mathbf{H}^T\mathbf{C}_\mathbf{e}^{-1}\mathbf{H}\right)^{-1} + \left(\mathbf{H}^T\mathbf{C}_\mathbf{e}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}_\mathbf{e}^{-1}\mathbf{C}_{\Delta\mathbf{f}}\mathbf{C}_\mathbf{e}^{-1}\mathbf{H}\left(\mathbf{H}^T\mathbf{C}_\mathbf{e}^{-1}\mathbf{H}\right)^{-1} \\ &= \mathbf{X}^{-1} + \mathbf{X}^{-1}\mathbf{Y}\mathbf{Z}\mathbf{Y}^T\mathbf{X}^{-1}\end{aligned} \tag{5-15}$$

where

$$\begin{aligned}\mathbf{X} &\triangleq \mathbf{H}^T\mathbf{C}_\mathbf{e}^{-1}\mathbf{H} \\ \mathbf{Y} &\triangleq \mathbf{H}^T\mathbf{C}_\mathbf{e}^{-1} \\ \mathbf{Z} &\triangleq \mathbf{K}\mathbf{C}_\mathbf{N}\mathbf{K}^T\end{aligned} \tag{5-16}$$

Compare with the case without sensor uncertainty, we can see the second term in $\text{var}\{\hat{\mathbf{p}}_e - \mathbf{p}_e\}$ is the increased variance. The trace of the second term of (5-15) is the decrease of accuracy due to the sensor states uncertainty.

From the above discussion, in the second stage estimation, if the sensor local state errors do not take into account, the effected item is the Jacobian matrix $\mathbf{H}$. That is the system matrix used in the iterative least square estimation for the second stage processing contains errors due to the navigation data uncertainty. A generalized least square, called Total Least Square (TLS) is usually used to solve this kind of problems. In appendix-4A, we introduce this method and find out whether it will not be applied to our problem. And we develop a new solution to solve this system uncertainty problem later.

5.2 CRLB Revisit

We know if there is no sensor navigation data uncertainty, Fisher information matrix of emitter location is calculated as

$$\mathrm{FIM}(\mathbf{p}_e)=\left[\frac{\partial \mathbf{f}(\mathbf{p}_e,\mathbf{B})}{\partial \mathbf{p}_e}\right]^T \mathbf{C}_\mathbf{e}^{-1}\left[\frac{\partial \mathbf{f}(\mathbf{p}_e,\mathbf{B})}{\partial \mathbf{p}_e}\right]\Bigg|_{\mathbf{p}_e}=\mathbf{H}^T\mathbf{C}_\mathbf{e}^{-1}\mathbf{H} \tag{5-17}$$

where all the matrices are defined previously. But since we do not know $\mathbf{B}$ exactly, the uncertainty of $\mathbf{B}$ will be counted into the CRLB of $\mathbf{p}_e$. So we need to re-evaluate the FIM of $\mathbf{p}_e$ based on the evaluation of FIM of $\left[\mathbf{p}_e^T \quad \mathbf{B}^T\right]^T$. There are two signal models related with $\mathbf{p}_e$ and $\mathbf{B}$ as

$$\begin{aligned}\hat{\mathbf{m}}&=\mathbf{f}(\mathbf{p}_e,\mathbf{B})+\mathbf{e}\\ \hat{\mathbf{B}}&=\mathbf{B}+\mathbf{N}\end{aligned} \tag{5-18}$$

The measurement vector is $\boldsymbol{\alpha}=\left[\hat{\mathbf{m}}^T \quad \hat{\mathbf{B}}^T\right]^T$. And we assume $\mathbf{N}$ is AWGN with zero mean and covariance matrix $\mathbf{C}_\mathbf{N}$; $\mathbf{N}$ and $\mathbf{e}$ are independent. Then the PDF of $\boldsymbol{\alpha}$ is given by

$$\begin{aligned}&\ln p\left(\left[\hat{\mathbf{m}}^T \quad \hat{\mathbf{B}}^T\right]^T;\left[\mathbf{m}^T \quad \mathbf{B}^T\right]^T\right)\\ &=\ln p\left(\hat{\mathbf{m}};\left[\mathbf{p}_e^T \quad \mathbf{B}^T\right]^T\right)+\ln p\left(\hat{\mathbf{B}};\mathbf{B}\right)\\ &=\alpha_1-\frac{1}{2}\left[\hat{\mathbf{m}}-\mathbf{f}(\mathbf{p}_e,\mathbf{B})\right]^T\mathbf{C}_\mathbf{e}^{-1}\left[\hat{\mathbf{m}}-\mathbf{f}(\mathbf{p}_e,\mathbf{B})\right]+\alpha_2-\frac{1}{2}\left(\hat{\mathbf{B}}-\mathbf{B}\right)^T\mathbf{C}_\mathbf{N}^{-1}\left(\hat{\mathbf{B}}-\mathbf{B}\right)\end{aligned} \tag{5-19}$$

where (if assume *N* sensors, *M* pairs and two-dimensional case)

$$\alpha_1=\frac{1}{(2\pi)^{2M}\left|\mathbf{C}_\mathbf{e}\right|^{1/2}} \text{ and } \alpha_2=\frac{1}{(2\pi)^{4N}\left|\mathbf{C}_\mathbf{N}\right|^{1/2}} \tag{5-20}$$

The FIM of $\left[\mathbf{p}_e^T \quad \mathbf{B}^T\right]^T$ will have the block structure as

$$\mathrm{FIM}\left(\left[\mathbf{p}_e^T,\mathbf{B}^T\right]^T\right)=\begin{bmatrix}\mathrm{FIM}(\mathbf{p}_e) & \mathrm{Cross\_FIM}(\mathbf{p}_e,\mathbf{B})\\ \mathrm{Cross\_FIM}^T(\mathbf{p}_e,\mathbf{B}) & \mathrm{FIM}(\mathbf{B})\end{bmatrix} \tag{5-21}$$

then we can evaluate each block individually.

Apply the FIM of AWGN case on (5-19), get

$$\mathrm{FIM}(\mathbf{p}_e) = \left[\frac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{p}_e}\right]^T \mathbf{C}_\mathbf{e}^{-1} \left[\frac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{p}_e}\right] = \mathbf{H}^T \mathbf{C}_\mathbf{e}^{-1} \mathbf{H} \tag{5-22}$$

and

$$\mathrm{FIM}(\mathbf{B}) = \left[\frac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{B}}\right]^T \mathbf{C}_\mathbf{e}^{-1} \left[\frac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{B}}\right] + \mathbf{C}_\mathbf{N}^{-1} = \mathbf{K}^T \mathbf{C}_\mathbf{e}^{-1} \mathbf{K} + \mathbf{C}_\mathbf{N}^{-1} \tag{5-23}$$

and

$$\mathrm{Cross\_FIM}(\mathbf{p}_e, \mathbf{B}) = \left[\frac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{p}_e}\right]^T \mathbf{C}_\mathbf{e}^{-1} \left[\frac{\partial \mathbf{f}(\mathbf{p}_e, \mathbf{B})}{\partial \mathbf{B}}\right] = \mathbf{H}^T \mathbf{C}_\mathbf{e}^{-1} \mathbf{K} \tag{5-24}$$

The CRLB matrix of $\left[\mathbf{p}_e^T \quad \mathbf{B}^T\right]^T$ will also have the block structure as

$$\begin{aligned} \mathrm{CRLB}\left(\left[\mathbf{p}_e^T, \mathbf{B}^T\right]^T\right) &= \left[\mathrm{FIM}\left(\left[\mathbf{p}_e^T, \mathbf{B}^T\right]^T\right)\right]^{-1} \\ &= \begin{bmatrix} \mathrm{CRLB}(\mathbf{p}_e) & \mathrm{Cross\_Covariance}(\mathbf{p}_e, \mathbf{B}) \\ \mathrm{Cross\_Covariance}^T(\mathbf{p}_e, \mathbf{B}) & \mathrm{CRLB}(\mathbf{B}) \end{bmatrix} \end{aligned} \tag{5-25}$$

Substitute (5-22), (5-23) and (5-24) into (5-21), and the matrix theory gives the CRLB of $\mathbf{p}_e$ as

$$\mathrm{CRLB}(\mathbf{p}_e) = \mathbf{X}^{-1} + \mathbf{X}^{-1}\mathbf{Y}\left(\mathbf{Z} - \mathbf{Y}^T\mathbf{X}^{-1}\mathbf{Y}\right)^{-1}\mathbf{Y}^T\mathbf{X}^{-1} \tag{5-26}$$

where $\mathbf{X}$, $\mathbf{Y}$ and $\mathbf{Z}$ have the same definition as (5-16).

The above analysis indicates that when there is sensor local state uncertainty, the CRLB of $\mathbf{p}_e$ will have one more term compare with the case without uncertainty. If we do not consider this uncertainty in our second stage estimation, the estimation covariance matrix can not reach the CRLB.

5.3 Updated Weighted Matrix

When we do not consider about the sensors' navigation data uncertainty, the weighted matrix is only related with the TDOA/FDOA measurement errors. And the estimation variance is (5-15). Ignore the linearization error $\Delta\mathbf{f}_H$, then (5-12) becomes

$$\hat{\mathbf{p}}_e - \mathbf{p}_e = \left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\left(\mathbf{e}+\mathbf{K}\mathbf{N}\right) \tag{5-27}$$

Update the weighted matrix $\mathbf{W}$ to

$$\mathbf{W}^{-1} = \mathrm{var}\left\{\mathbf{e}+\mathbf{K}\mathbf{N}\right\} = \mathbf{C}_\mathbf{e} + \mathbf{K}\mathbf{C}_\mathbf{N}\mathbf{K}^T \tag{5-28}$$

then

$$\begin{aligned}\mathrm{var}\left\{\hat{\mathbf{p}}_e - \mathbf{p}_e\right\} &= \left[\mathbf{H}^T\left(\mathbf{C}_\mathbf{e} + \mathbf{K}\mathbf{C}_\mathbf{N}\mathbf{K}^T\right)^{-1}\mathbf{H}\right]^{-1} \\ &= \mathbf{X}^{-1} + \mathbf{X}^{-1}\mathbf{Y}\left(\mathbf{Z} - \mathbf{Y}^T\mathbf{X}^{-1}\mathbf{Y}\right)^{-1}\mathbf{Y}^T\mathbf{X}^{-1}\end{aligned} \tag{5-29}$$

where $\mathbf{X}$ $\mathbf{Y}$ and $\mathbf{Z}$ are defined as (5-16).

The trace of the second term of (5-29) comparing to the trace of the second term of (5-15) is the increased accuracy for geo-location estimation, if we use $\mathbf{W}^{-1} = \mathbf{C}_\mathbf{e} + \mathbf{K}\mathbf{C}_\mathbf{N}\mathbf{K}^T$ instead of $\mathbf{W}^{-1} = \mathbf{C}_\mathbf{e}$.

And also when the navigation date error vector is AWGN and $\hat{\mathbf{B}}$ is close enough to $\mathbf{B}$, if we use (5-28) as the weighted matrix, the covariance matrix of $\mathbf{p}_e$ can reach the CRLB of $\mathbf{p}_e$ as in (5-26).

5.4 Estimation Accuracy with Local State Error

In the previous session, we developed the updated CRLB of geo-location under the case with sensors' states uncertainty and give a solution to reach it. In chapter 3, we have

discussed all the related items of emitter location accuracy, but we do not develop more details about to what extent sensor local states estimation errors will contribute to the estimation accuracy. In this section, we analyze the sensitivity of the estimation accuracy with respect to local state. From (3-60), the uncertainty caused by sensors navigation data is as

$$\Sigma_m = \left(\frac{\partial f_m}{\partial \boldsymbol{\beta}_m} \right)^T d\boldsymbol{\beta}_m \tag{5-30}$$

And for notation simplicity, we discuss one pair only and assume the pair is paired by sensors S_1 and S_2. All the typical values of sensors' navigation data are from ?????

5.4.1 TDOA

The TDOA signal model is influenced by sensor position vector $\boldsymbol{\beta}_m = \left[\mathbf{s}_1^T \quad \mathbf{s}_2^T \right]^T$ only.

$$\Sigma_m = \left(\frac{\partial \tau_m}{\partial \boldsymbol{\beta}_m} \right)^T \Delta\boldsymbol{\beta}_m = \left\| \frac{\partial \tau_m}{\partial \boldsymbol{\beta}_m} \right\| \cdot \left\| \Delta\boldsymbol{\beta}_m^{\perp} \right\| \tag{5-31}$$

where $\partial \tau_m / \partial \boldsymbol{\beta}_m$ is the gradient vector of τ_m with respect to $\boldsymbol{\beta}_m$. And $\Delta\boldsymbol{\beta}_m^{\perp}$ is the component of $\Delta\boldsymbol{\beta}_m$ in the direction of $\partial \tau_m / \partial \boldsymbol{\beta}_m$, which is the normal to the surface defined by τ_m. Then

$$\sigma_{\Sigma_m}^2 = \left\| \frac{\partial \tau_m}{\partial \boldsymbol{\beta}_m} \right\|^2 \cdot \left\| \Delta\boldsymbol{\beta}_m^{\perp} \right\|^2 \tag{5-32}$$

The uncertainty in $\boldsymbol{\beta}_m$ is reflected in τ_m through

$$\left\| \frac{\partial \tau_m}{\partial \boldsymbol{\beta}_m} \right\| = \frac{1}{c} \left\| \begin{bmatrix} -\mathbf{u}_1 \\ \mathbf{u}_2 \end{bmatrix} \right\| = \frac{\sqrt{2}}{c} \tag{5-33}$$

And the typical error of position of the on-board navigation system is $\left\|\Delta \mathbf{s}_m\right\|$ is 10*meter*.

Then

$$\partial_{\Sigma_m} \approx \frac{\sqrt{2}}{c} \cdot 10\sqrt{2} = \frac{20}{c} \tag{5-34}$$

5.4.2 FDOA

The FDOA signal model is influenced by both sensor position vector $\boldsymbol{\beta}_m = \begin{bmatrix} \mathbf{s}_1^T & \mathbf{s}_2^T \end{bmatrix}^T$ and velocity vector $\dot{\boldsymbol{\beta}}_m = \begin{bmatrix} \dot{\mathbf{s}}_1^T & \dot{\mathbf{s}}_2^T \end{bmatrix}^T$. And can be evaluated from

$$\partial_{\Sigma_m}^2 = \left\| \frac{\partial \upsilon_m}{\partial \boldsymbol{\beta}_m} \right\|^2 \cdot \left\| \Delta \boldsymbol{\beta}_m^{\perp} \right\|^2 + \left\| \frac{\partial \upsilon_m}{\partial \dot{\boldsymbol{\beta}}_m} \right\|^2 \cdot \left\| \Delta \dot{\boldsymbol{\beta}}_m^{\perp} \right\|^2 \tag{5-35}$$

where $\partial \upsilon_m / \partial \boldsymbol{\beta}_m$, $\partial \upsilon_m / \partial \dot{\boldsymbol{\beta}}_m$ are the gradient vectors of υ_m with respect to $\boldsymbol{\beta}_m$, $\dot{\boldsymbol{\beta}}_m$ respectively. The uncertainty in $\boldsymbol{\beta}_m$ is reflected in υ_m through

$$\left\| \frac{\partial \upsilon_m}{\partial \boldsymbol{\beta}_m} \right\| = \frac{f_e}{c} \left\| \begin{bmatrix} \dfrac{\dot{\mathbf{s}}_1 - \mathbf{u}_1^T \dot{\mathbf{s}}_1 \mathbf{u}_1}{r_1} \\ \dfrac{\dot{\mathbf{s}}_2 - \mathbf{u}_2^T \dot{\mathbf{s}}_2 \mathbf{u}_2}{r_2} \end{bmatrix} \right\| \le \frac{f_e}{c} \left(\frac{\left\| \dot{\mathbf{s}}_1 \right\|}{r_1} + \frac{\left\| \dot{\mathbf{s}}_2 \right\|}{r_2} \right) \tag{5-36}$$

For the typical values of r_k and $\| \dot{\mathbf{s}}_k \|$, which are 50*km* and 250*m/s*, $\| \dot{\mathbf{s}}_1 \| / r_1 + \| \dot{\mathbf{s}}_2 \| / r_2$ is on the order of $10^{-2}\sqrt{2}/s$.

The uncertainty in $\dot{\boldsymbol{\beta}}_m$ is reflected in υ_m through

$$\left\| \frac{\partial \upsilon_m}{\partial \dot{\boldsymbol{\beta}}_m} \right\| = \frac{f_e}{c} \left\| \begin{bmatrix} \mathbf{u}_1 \\ -\mathbf{u}_2 \end{bmatrix} \right\| = \sqrt{2} \frac{f_e}{c} \tag{5-37}$$

The typical error of velocity of the on-board navigation system is $\left\|\Delta \dot{\mathbf{s}}_k\right\|$ is 0.02*m/s*.

Therefore

$$\partial_{\Sigma_m} \approx \frac{f_e}{c}\sqrt{2}(10^{-2}\cdot 10 + 0.02) \approx \frac{f_e}{c}\sqrt{2}(0.1 + 0.02) \qquad (5\text{-}38)$$

Compare (5-34) and (5-38) reveals that the influence of uncertainty of position on TDOA can be ignored comparing with the effects of uncertainty of position and velocity on FDOA.

5.5 Appendix-5A Total Least Square Method for NAV Data Errors

5A.1 Introduction of TLS

- **Least Squares**

In order to introduce TLS, we briefly talk about some basic points about LS. Assuming signal has linear model on the parameter vector $\boldsymbol{\theta}$ as

$$\mathbf{s} = \mathbf{H}\boldsymbol{\theta} \qquad (5\text{-}39)$$

where $\mathbf{H}$ is a known $N \times p$ matrix $(N > p)$ of full rank. The matrix $\mathbf{H}$ is refereed to as the observation/system matrix. Knowing $\mathbf{H}$ and $\mathbf{s}$ we should be able to find the parameter vector $\boldsymbol{\theta}$ that made the observable $\mathbf{s}$, since $\mathbf{s} \in \text{range}(\mathbf{H})$. However, in practice we often only have imperfect knowledge of $\mathbf{s}$ due to the measurement error, which is $\tilde{\mathbf{s}} = \mathbf{s} + \Delta\mathbf{s}$. Then we are going to solve $\tilde{\mathbf{s}} = \mathbf{H}\boldsymbol{\theta}$. There is no solution for this equation, since $\tilde{\mathbf{s}} \notin \text{range}(\mathbf{H})$. Then the problem is: Given the matrix $\mathbf{H}$ and the imperfect $\mathbf{s}$, find an appropriate estimate of $\boldsymbol{\theta}$. The LS approach finds an estimate $\hat{\boldsymbol{\theta}}$ such that the residual defined as $\mathbf{r} = \mathbf{H}\hat{\boldsymbol{\theta}} - \tilde{\mathbf{s}}$ has minimum norm; that is, $\left\|\mathbf{H}\hat{\boldsymbol{\theta}} - \tilde{\mathbf{s}}\right\|_2$ is minimized. In other words:

Least Square: Minimize $\|\mathbf{r}\|_2$ subject to $\tilde{\mathbf{s}} + \mathbf{r} \in \text{range}(\mathbf{H})$ and then find the $\hat{\boldsymbol{\theta}}$ that satisfies $\tilde{\mathbf{s}} + \mathbf{r} = \mathbf{H}\hat{\boldsymbol{\theta}}$.

- **Total Least Squares**

The signal model is still as (5-39), but $\mathbf{H}$ is perturbed by system errors to be $\tilde{\mathbf{H}} = \mathbf{H} + \Delta\mathbf{H}$. We are going to solve $\tilde{\mathbf{s}} = \tilde{\mathbf{H}}\boldsymbol{\theta}$. There is no solution for this equation either, since $\tilde{\mathbf{s}} \notin \text{range}(\tilde{\mathbf{H}})$. Recall that for standard LS we sought to find a minimum norm vector $\mathbf{r}$ that would bring $\tilde{\mathbf{s}}$ back into $\text{range}(\mathbf{H})$. Here we seek simultaneously to find a vector $\mathbf{r}$ and a matrix $\mathbf{E}$ such that

$$\tilde{\mathbf{s}} + \mathbf{r} \in \text{range}(\tilde{\mathbf{H}} + \mathbf{E}) \tag{5-40}$$

Then the total least square can be viewed as:

Total Least Square: Find the smallest perturbations $\mathbf{r}$ and $\mathbf{E}$ such that $\tilde{\mathbf{s}} + \mathbf{r} \in \text{range}(\tilde{\mathbf{H}} + \mathbf{E})$ and then find the $\hat{\boldsymbol{\theta}}$ that satisfies $\tilde{\mathbf{s}} + \mathbf{r} = (\tilde{\mathbf{H}} + \mathbf{E})\hat{\boldsymbol{\theta}}$.

The criterion taken in TLS to measure the joint size of perturbations $\mathbf{r}$ and $\mathbf{E}$ is to form the $N \times (p+1)$ matrix $\boldsymbol{\Delta} = [\mathbf{E} \mid \mathbf{r}]$, and measure the size of $\boldsymbol{\Delta}$ in some sense. The norm of a matrix is used for this purpose. There are many matrix norms that can be defined, but the one used here for TLS is the Fresenius norm that

$$\|\boldsymbol{\Delta}\|_F = \sqrt{\sum_i \sum_j \Delta_{ij}^2} \tag{5-41}$$

Then the definition of TLS can be rewritten as:

Total Least Square: Minimize $\|[\mathbf{E} \mid \mathbf{r}]\|_F$ subject to $\tilde{\mathbf{s}} + \mathbf{r} \in \text{range}(\tilde{\mathbf{H}} + \mathbf{E})$ and then find the $\hat{\boldsymbol{\theta}}$ that satisfies $\tilde{\mathbf{s}} + \mathbf{r} = (\tilde{\mathbf{H}} + \mathbf{E})\hat{\boldsymbol{\theta}}$.

The equation to solve is

$$\tilde{\mathbf{s}} + \mathbf{r} = (\tilde{\mathbf{H}} + \mathbf{E})\boldsymbol{\theta} \tag{5-42}$$

which can be rewritten as

$$(\mathbf{A} + \boldsymbol{\Delta})\begin{bmatrix} \boldsymbol{\theta} \\ -\mathbf{1} \end{bmatrix} = \mathbf{0} \tag{5-43}$$

where $\mathbf{A} = \left[\tilde{\mathbf{H}} \mid \tilde{\mathbf{s}}\right]$. To get a nontrivial solution to (5-43) requires that $\boldsymbol{\Delta}$ be such that $\mathbf{A} + \boldsymbol{\Delta}$ is rank deficient. That is to find $\tilde{\mathbf{A}} = \mathbf{A} + \boldsymbol{\Delta}$ such that $\tilde{\mathbf{A}}$ is the closest rank-deficient matrix to $\mathbf{A}$. Then the definition of TLS can be rewritten as:

Total Least Square: Find the smallest perturbation $\boldsymbol{\Delta}$ that makes $\tilde{\mathbf{A}} = \left[\tilde{\mathbf{H}} \mid \tilde{\mathbf{s}}\right] + \boldsymbol{\Delta}$ rank deficient then find the $\hat{\boldsymbol{\theta}}$ such that $\boldsymbol{\beta} = [\hat{\boldsymbol{\theta}}^T, -\mathbf{1}^T]^T$ is in the null space of $\tilde{\mathbf{A}}$.

Matrix theory says that such a matrix $\tilde{\mathbf{A}}$ is found by replacing the smallest singular value of $\mathbf{A}$ by zero. Thus, the vector that solves $\tilde{\mathbf{A}}\boldsymbol{\beta} = \mathbf{0}$ must be orthogonal to the singular vectors of $\mathbf{A}$ that remain in $\tilde{\mathbf{A}}$. Since the singular vector $\mathbf{v}$ removed from $\mathbf{A}$ to make $\tilde{\mathbf{A}}$ is orthogonal to the singular vectors that remain in $\tilde{\mathbf{A}}$, it must satisfy $\tilde{\mathbf{A}}\mathbf{v} = \mathbf{0}$. Thus, $\boldsymbol{\beta}$ is proportional to $\mathbf{v}$, and all that is needed is to make the last entry in $\mathbf{v}$ be -1 to get the form needed for $\boldsymbol{\beta}$. So if we take $\mathbf{v}$ and divide all of its elements by the negative of its last element we get $\boldsymbol{\beta}$.

We know the weighted LS (WLS) is used to normalize the residue. So correspondingly there is a WTLS method. Let $\mathbf{W}$ be a given $N \times N$ left-side weighting matrix and $\mathbf{T}$ be a given $(p+1)\times(p+1)$ right-side weighting matrix. Then the WTLS becomes to:

Weighted TLS: Minimize $\|\mathbf{W}[\mathbf{E} \mid \mathbf{r}]\mathbf{T}\|_F$ subject to $\tilde{\mathbf{s}} + \mathbf{r} \in \text{range}(\tilde{\mathbf{H}} + \mathbf{E})$ and then find the $\hat{\boldsymbol{\theta}}$ that satisfies $\tilde{\mathbf{s}} + \mathbf{r} = (\tilde{\mathbf{H}} + \mathbf{E})\hat{\boldsymbol{\theta}}$.

- **Performance of Total Least-Squares**

The problem to be solved is $\tilde{\mathbf{s}} = \tilde{\mathbf{H}}\boldsymbol{\theta}$, where $\tilde{\mathbf{H}}_{N\times p}$ and $\tilde{\mathbf{s}}_{N\times 1}$ are perturbed versions of some $\mathbf{H}$ and $\mathbf{s}$. We will assume that the perturbations have uniform levels in all columns of $\tilde{\mathbf{H}}$ and $\tilde{\mathbf{s}}$; this uniformity may have been achieved through the weighting matrices $\mathbf{W}$ and $\mathbf{T}$. We also assume that $\tilde{\mathbf{H}}$ has full rank. For notational purposes, define the following singular value decompositions:

$$\begin{aligned} \mathbf{H} &= \sum_{i=1}^{p} \sigma_i' \mathbf{u}_i' \mathbf{v}_i'^T \qquad \tilde{\mathbf{H}} = \sum_{i=1}^{p} \tilde{\sigma}_i' \tilde{\mathbf{u}}_i' \tilde{\mathbf{v}}_i'^T \\ [\mathbf{H} \mid \mathbf{s}] &= \sum_{i=1}^{p+1} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad [\tilde{\mathbf{H}} \mid \tilde{\mathbf{s}}] = \sum_{i=1}^{p+1} \tilde{\sigma}_i \tilde{\mathbf{u}}_i \tilde{\mathbf{v}}_i^T \end{aligned} \tag{5-44}$$

The performance advantage of TLS compared to LS has been studied and it has been shown that, for the assumptions made above, TLS performs better than LS when

$$\left[\frac{\|\Delta\mathbf{H}\|_2}{\|[\Delta\mathbf{H} \mid \Delta\mathbf{s}]\|_2}\right] \cdot \left[\frac{\tilde{\sigma}_p}{\tilde{\sigma}_p'}\right] > 1 \tag{5-45}$$

where $\|\mathbf{M}\|_2$ denotes the two-norm of matrix $\mathbf{M}$ defined as

$$\|\mathbf{M}\|_2 = \sup_{\|x\|_2=1} \|\mathbf{M}\mathbf{x}\|_2 \tag{5-46}$$

TLS performance improves as the quantity on the left-hand side of (5-45) increases. Because we assume uniform level of perturbations, the ratio of norms in the first term in (5-45) is

$$\left[\frac{\|\Delta\mathbf{H}\|_2}{\|[\Delta\mathbf{H} \mid \Delta\mathbf{s}]\|_2}\right] \approx 1 \tag{5-47}$$

Thus, we only need to look at the ratio of singular values in (5-45), which depends on the relative structures of $\tilde{\mathbf{H}}$ and $[\tilde{\mathbf{H}} \mid \tilde{\mathbf{s}}]$. It is shown in [] that the ratio $\tilde{\sigma}_n / \tilde{\sigma}_n'$ is

influenced by (i) the norm of vector $\mathbf{s}$ and (ii) the angle between the vector $\mathbf{s}$ and the n^{th} left singular vector $\mathbf{u}'_n$ of $\mathbf{H}$. To see this, first note that because the perturbations are at uniform levels, we can investigate this phenomenon using the unperturbed matrices and their ratio σ_n / σ'_n; this is because the uniform-level perturbations will uniformly impact the numerator and denominator of σ_n / σ'_n to create $\tilde{\sigma}_n / \tilde{\sigma}'_n$. So from (5-44) we see that

$$\begin{aligned} \mathbf{H}^T \mathbf{u}'_n &= \sigma'_n \mathbf{v}'_n \\ \begin{bmatrix} \mathbf{H}^T \\ \mathbf{s}^T \end{bmatrix} \mathbf{u}_n &= \begin{bmatrix} \mathbf{H}^T \mathbf{u}_n \\ \mathbf{s}^T \mathbf{u}_n \end{bmatrix} = \sigma_n \mathbf{v}_n \end{aligned} \tag{5-48}$$

When $\mathbf{H}^T \mathbf{u}_n$ and $\mathbf{s}^T \mathbf{u}_n$ simultaneously become large, $\left[\mathbf{H} \mid \mathbf{s}\right]^T \mathbf{u}_n$ will have maximum length. When $\mathbf{u}_n = \mathbf{u}'_n$ the length of $\mathbf{H}^T \mathbf{u}_n$ is maximized and then if in addition $\mathbf{s}$ is aligned with $\mathbf{u}_n = \mathbf{u}'_n$, the magnitude of $\mathbf{s}^T \mathbf{u}_n$ is maximized. Furthermore, the larger the norm of $\mathbf{s}$, the larger the increase in σ_n will be. Thus, TLS will perform at its best with respect to LS when $\mathbf{s}$ is aligned with the n^{th} left singular vector $\mathbf{u}'_n$ of $\mathbf{H}$ and has a large norm.

5A.2 TLS Performance for Emitter Location Estimation

The signal model for TDOA/FDOA emitter location problem is

$$\mathbf{m} = \mathbf{f}(\mathbf{p}_e, \mathbf{B}) \tag{5-49}$$

and the observed vector model is

$$\hat{\mathbf{m}} = \mathbf{f}(\mathbf{p}_e, \hat{\mathbf{B}}) + \mathbf{e} \tag{5-50}$$

Vectors are defined as in (5-3). We are going to find the solution of the linearized (5-50) as

$$\hat{\mathbf{m}}_1 = \mathbf{H}_{\hat{\mathbf{B}}} \cdot \mathbf{p}_e + \mathbf{e} \tag{5-51}$$

where

$$\begin{aligned} \hat{\mathbf{m}}_1 &= \hat{\mathbf{m}} - \mathbf{f}(\mathbf{p}_e^0, \hat{\mathbf{B}}) + \mathbf{H}_{\hat{\mathbf{B}}} \mathbf{p}_e^0 \\ \mathbf{H}_{\hat{\mathbf{B}}} &= \left. \frac{\partial \mathbf{f}(\mathbf{p}_e, \hat{\mathbf{B}})}{\partial \mathbf{p}_e} \right|_{\mathbf{p}_e = \mathbf{p}_e^0} \end{aligned} \tag{5-52}$$

where $\mathbf{p}_e^0$ is the normal value of $\mathbf{p}_e$. (5-51) is viewed as a perturbed version of

$$\mathbf{m}_1 = \mathbf{H}_{\Theta} \cdot \mathbf{p}_e \tag{5-53}$$

where $\mathbf{m}_1 = \mathbf{m} - \mathbf{f}(\mathbf{p}_e^0, \mathbf{B}) + \mathbf{H}_{\mathbf{B}} \mathbf{p}_e^0$. And if we assume $\mathbf{p}_e^0$ is close enough to $\mathbf{p}_e$, then $\mathbf{m}_1 \approx \mathbf{0}$. (5-53) becomes a homogenous linear system.

For the two-dimensional system we used, $\mathbf{H}$ has two columns as

$$\begin{aligned} \mathbf{h}_x &= \begin{bmatrix} \dfrac{x_1 - x_e}{r_1} - \dfrac{x_2 - x_e}{r_2} \\ \left(\dfrac{\mathbf{u}_1^T \dot{\mathbf{s}}_1 (x_1 - x_e)}{r_1^2} - \dfrac{\dot{x}_1}{r_1} \right) - \left(\dfrac{\mathbf{u}_2^T \dot{\mathbf{s}}_2 (x_2 - x_e)}{r_2^2} - \dfrac{\dot{x}_2}{r_2} \right) \end{bmatrix} \\ \mathbf{h}_y &= \begin{bmatrix} \dfrac{y_1 - y_e}{r_1} - \dfrac{y_2 - y_e}{r_2} \\ \left(\dfrac{\mathbf{u}_1^T \dot{\mathbf{s}}_1 (y_1 - y_e)}{r_1^2} - \dfrac{\dot{y}_1}{r_1} \right) - \left(\dfrac{\mathbf{u}_2^T \dot{\mathbf{s}}_2 (y_2 - y_e)}{r_2^2} - \dfrac{\dot{y}_2}{r_2} \right) \end{bmatrix} \end{aligned} \tag{5-54}$$

Since $\mathbf{m}_1 = \mathbf{0}$, apply this special case on (5-44)

$$\begin{aligned} \mathbf{H}_{\mathbf{B}} &= \sum_{i=1}^{2} \sigma_i' \mathbf{u}_i' \mathbf{v}_i'^T \qquad \mathbf{H}_{\hat{\mathbf{B}}} = \sum_{i=1}^{2} \tilde{\sigma}_i' \tilde{\mathbf{u}}_i' \tilde{\mathbf{v}}_i'^T \\ [\mathbf{H}_{\mathbf{B}} \mid \mathbf{0}] &= \sum_{i=1}^{3} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad [\mathbf{H}_{\hat{\mathbf{B}}} \mid \mathbf{e}] = \sum_{i=1}^{3} \tilde{\sigma}_i \tilde{\mathbf{u}}_i \tilde{\mathbf{v}}_i^T \end{aligned} \tag{5-55}$$

Then we need to look at the ratio of the 2^{nd} singular value of $[\mathbf{H}_{\hat{\mathbf{B}}} \mid \mathbf{e}]$ to the 2^{nd} singular value of $\mathbf{H}_{\hat{\mathbf{B}}}$. Since the navigation data is obtained by some combined methods to make it close enough to the true values, then the perturbations are small on $\mathbf{H}$. we can

instead just look at the singular values of the unperturbed matrices because the perturbations will have little effect in making this ratio much different from the ratio of singular values of the perturbed case. The singular values of the unperturbed matrices of interest here are

$$\begin{aligned} \mathbf{H}_{\hat{\mathbf{B}}}: &\quad \sigma_1', \sigma_2' \\ \left[\mathbf{H}_{\hat{\mathbf{B}}} \,|\, \mathbf{0}\right]: &\quad \sigma_1', \sigma_2', 0 \end{aligned} \tag{5-56}$$

Thus, the perturbed ratio approximately equals the unperturbed ratio, which equals 1; therefore the performance of TLS is equivalent to that of LS for this case. It is the fact that the underlying unperturbed problem to be solved is a homogeneous equation that makes TLS ineffective compared to LS. These insights are consistent with the general observations in: TLS has its greatest advantage when $\mathbf{s}$ has a large length and is aligned in direction with the n^{th} left singular vector $\mathbf{u}_n'$ of $\mathbf{H}$; however, for the homogeneous case $\mathbf{s}$ has zero length and is orthogonal to $\mathbf{u}_n'$ and therefore provides no advantage for TLS with respect to LS.

Our conclusion is that since the sensor local states uncertainty is small, TLS has no advantage than LS in estimating emitter location. We will not use TLS to improve the accuracy.

6 More Issues about Estimation Accuracy

From above discussion, the accuracy of the emitter location estimation is depending on three aspects: (i) Accuracy of the first stage TDOA/FDOA measurements; (ii) The current sensors' navigation data; and (iii) The sensors' local states accuracy. For the first aspect, one important thing we need to consider is that whether the estimates are obtained from centralized maximum likelihood method or de-centralized maximum likelihood method, since the covariance matrices maybe different. This issue has been talked before, but the result as in [14] was only for acoustic scenario as we discussed in chapter 3. In this chapter, we develop the result for RF scenario. For the second aspect, we can try to find the next optimal states based on the current measurements and next researchable set. And enable the sensor to be as error-insensitive as possible to the various sources of error. We have discussed the third aspect and developing some results and updated algorithms to increase the estimation accuracy with sensors' states uncertainty.

6.1 Gaussian Maximum Likelihood Estimator

In practice, the sensor network type-III is widely be used. There is a center sensor which has all the sensors' navigation data and whole/part received data. This sensor has super computation ability than other sensors. And it can control other sensors, such as make the sensor selection and decide the next state for sensors. There are two methods to compute the TDOA/FDOA for all the pairs at the center sensor. One is centralized, which estimates all the TDOA/FDOA parameters together; the other one is de-centralized, which gets the TDOA/FDOA estimations pair-wised. The comparison of centralized

maximum likelihood method and de-centralized maximum likelihood method is as Figure 24, here the ML method used is cross-correlation method.

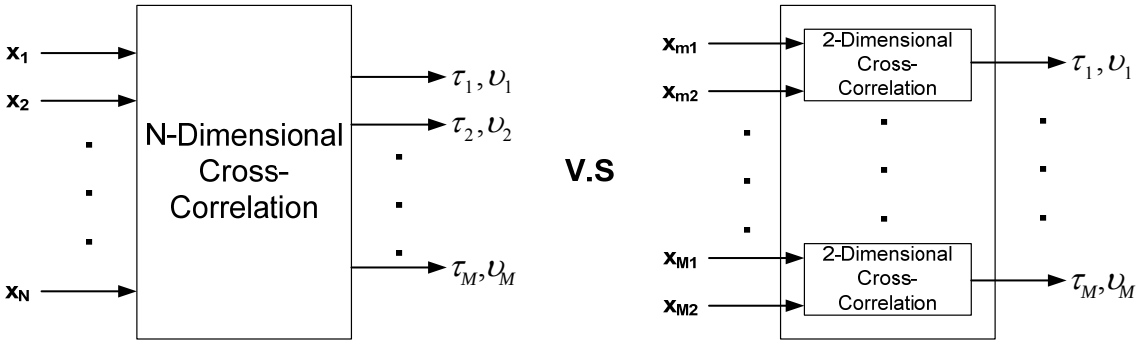

Figure 24 Centralized cross-correlation vs. de-centralize cross-correlation method

The centralized cross-correlation is a *N*-dimensional processing procedure which deals with the signal vector $\mathbf{x} = \begin{bmatrix} \mathbf{x}_1^T & \cdots & \mathbf{x}_N^T \end{bmatrix}^T$ together, $\mathbf{x}_n$ is data vector from each receiver. And the cross-correlator is a *N*-dimensional estimator to get the *M* pair estimates. On the other hand, the de-centralized one has *M* 2-dimentional cross-correlators. Each one only deals with data vector $\mathbf{x}_m = \begin{bmatrix} \mathbf{x}_{m,1}^T & \mathbf{x}_{m,2}^T \end{bmatrix}^T$ and gives the estimation for one pair only.

6.1.1 CML and DML Estimators

(1) CML Estimator Structure

Since we are dealing with the active radar signals (as we discussed in chapter 3), signal model is deterministic signal plus additive Gaussian noise as

$$x(t) = s(t;\boldsymbol{\theta}) + w(t) \tag{4-1}$$

$s(t;\boldsymbol{\theta})$ represents deterministic signal $s(t)$ parameterized by vector $\boldsymbol{\theta}$, which is a $p \times 1$ deterministic vector . Assume $w(t)$ is zero-mean Gaussian noise.

Based on the signal model, the probability density function of Gaussian data vector $\mathbf{x}$ (sampled data vector of $x(t)$) is:

$$p(\mathbf{x};\boldsymbol{\theta}) = \frac{1}{(2\pi)^{N/2}\det^{1/2}(\mathbf{C}_w)}\exp\left[-\frac{1}{2}(\mathbf{x}-\mathbf{s}_{\boldsymbol{\theta}})^T\mathbf{C}_w^{-1}(\mathbf{x}-\mathbf{s}_{\boldsymbol{\theta}})\right] \tag{4-2}$$

where $\mathbf{s}_{\boldsymbol{\theta}}$ is the signal vector, which depends on $\boldsymbol{\theta}$; $\mathbf{C}_w$ is noise covariance matrix.

The maximum likelihood estimate $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is any solution of the likelihood equations

$$z_i(\hat{\boldsymbol{\theta}};\mathbf{x}) = \frac{\partial}{\partial\theta_i}\ln p(\mathbf{x};\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}} = 0 \quad i = 1,2,...,p \tag{4-3}$$

Substitute (4-2) in (4-3) and get,

$$z_i(\hat{\boldsymbol{\theta}};\mathbf{x}) = \left(\frac{\partial\hat{\mathbf{s}}_{\boldsymbol{\theta}}}{\partial\theta_i}\right)^T\mathbf{C}_w^{-1}(\mathbf{x}-\hat{\mathbf{s}}_{\boldsymbol{\theta}}) = 0 \tag{4-4}$$

The hat '^' denotes the deterministic signal and its derivative with respect to θ_i are evaluated at $\boldsymbol{\theta} = \hat{\boldsymbol{\theta}}$. Thus

$$\frac{\partial\hat{\mathbf{s}}_{\boldsymbol{\theta}}}{\partial\theta_i} = \frac{\partial\mathbf{s}_{\boldsymbol{\theta}}}{\partial\theta_i}\Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \qquad \hat{\mathbf{s}}_{\boldsymbol{\theta}} = \mathbf{s}_{\boldsymbol{\theta}}|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \tag{4-5}$$

(2) DML Estimator Structure

We partition all the received data vector into M sub-vectors $\{\mathbf{x}^{(m)}\}_{m=1}^{M}$, which satisfied $\bigcup_{m=1}^{M}\mathbf{x}^{(m)} = \mathbf{x}$, and partition the parameter vector into M corresponding sub-vectors $\{\boldsymbol{\theta}^{(m)}\}_{m=1}^{M}$. (For example, in TDOA/FDOA case, if pair one is paired by sensor S_1 and S_2, then $\mathbf{x}^{(1)} = \left[\mathbf{x}_1^T, \mathbf{x}_2^T\right]^T$ and $\boldsymbol{\theta}^{(1)} = \left[\tau_1, \upsilon_1\right]^T$.) Then the mean vector for the sub-vector $\mathbf{x}^{(m)}$ depends on $\boldsymbol{\theta}^{(m)}$ only as

$$\mathrm{E}\{\mathbf{x}^{(m)}\} = \mathbf{s}_{\boldsymbol{\theta}^{(m)}}^{(m)} \tag{4-6}$$

where $\mathbf{s}_{\boldsymbol{\theta}^{(m)}}^{(m)}$ is simply a sub-vector of $\mathbf{s}_{\boldsymbol{\theta}}$.

Decomposition of the data vector and the parameter vector suggests a form of an estimator consisting of *M* ML data processors in parallel. The m^{th} set of likelihood functions is given by

$$\mathbf{z}^{(m)}\left(\hat{\boldsymbol{\theta}}^{(m)};\mathbf{x}^{(m)}\right)=\left(\frac{\partial\hat{\mathbf{s}}_{\boldsymbol{\theta}^{(m)}}^{(m)}}{\partial\boldsymbol{\theta}^{(m)}}\right)^{T}\left(\mathbf{C}_{w}^{(m)}\right)^{-1}\left(\mathbf{x}^{(m)}-\hat{\mathbf{s}}_{\boldsymbol{\theta}^{(m)}}^{(m)}\right) \tag{4-7}$$

(3) DML Estimator Performance

To evaluate the DML method, we need to calculate the bias and error covariance matrix of the estimates. Here we begin with linearize each sub-vector of likelihood function $\mathbf{z}^{(m)}$ with respect to $\hat{\boldsymbol{\theta}}^{(m)}$ using the following stochastic approximation

$$\mathbf{z}^{(m)}(\hat{\boldsymbol{\theta}}^{(m)})=\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})+\mathbf{J}(\boldsymbol{\theta}^{(m)})\left(\hat{\boldsymbol{\theta}}^{(m)}-\boldsymbol{\theta}^{(m)}\right) \tag{4-8}$$

where $\mathbf{J}(\boldsymbol{\theta}^{(m)})$ is the matrix of partial derivative with respect to $\hat{\boldsymbol{\theta}}^{(m)}$ and evaluated at $\boldsymbol{\theta}^{(m)}$, the $(i,j)^{th}$ element of it is

$$\mathbf{J}_{ij}(\boldsymbol{\theta}^{(m)})=-\mathrm{E}\left\{\frac{\partial z_{i}^{(m)}}{\partial\hat{\theta}_{j}^{(m)}}\right\}=-\mathrm{E}\left\{\frac{\partial^{2}\ln p(\mathbf{x}^{(m)};\boldsymbol{\theta}^{(m)})}{\partial\hat{\theta}_{i}^{(m)}\partial\hat{\theta}_{j}^{(m)}}\right\} \tag{4-9}$$

We can see $\mathbf{J}_{ij}(\boldsymbol{\theta}^{(m)})$ is exactly the $(i,j)^{th}$ element of the FIM of $\boldsymbol{\theta}^{(m)}$ computed from $p(\mathbf{r}^{(m)};\boldsymbol{\theta}^{(m)})$.

For ML method, $\hat{\boldsymbol{\theta}}^{(m)}$ is the solution of $\mathbf{z}^{(m)}(\hat{\boldsymbol{\theta}}^{(m)})=\mathbf{0}$. Then from (4-8) get

$$\hat{\boldsymbol{\theta}}^{(m)}=\boldsymbol{\theta}^{(m)}-\mathbf{J}^{-1}(\boldsymbol{\theta}^{(m)})\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)}) \tag{4-10}$$

Then

$$\mathrm{E}\left\{\hat{\boldsymbol{\theta}}^{(m)}\right\}=\boldsymbol{\theta}^{(m)}-\mathbf{J}^{-1}(\boldsymbol{\theta}^{(m)})\cdot\mathrm{E}\left\{\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})\right\} \tag{4-11}$$

and

$$\mathrm{var}\left\{\hat{\boldsymbol{\theta}}^{(m)}\right\} = \mathbf{J}^{-1}(\boldsymbol{\theta}^{(m)}) \cdot \mathrm{var}\left\{\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})\right\} \cdot \mathbf{J}^{-1}(\boldsymbol{\theta}^{(m)}) \tag{4-12}$$

Evaluate $\mathrm{E}\left\{\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})\right\}$ and $\mathrm{var}\left\{\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})\right\}$ by substituting (4-7) and (4-3) in then

$$\mathrm{E}\left\{\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})\right\} = \frac{\partial \mathbf{s}^{(m)}_{\boldsymbol{\theta}^{(m)}}}{\partial \boldsymbol{\theta}^{(m)}} (\mathbf{C}^{(m)}_w)^{-1} \left[\mathrm{E}\left\{\mathbf{x}^{(m)}\right\} - \mathbf{s}^{(m)}_{\boldsymbol{\theta}^{(m)}}\right] = \frac{\partial \mathbf{s}^{(m)}_{\boldsymbol{\theta}^{(m)}}}{\partial \boldsymbol{\theta}^{(m)}} (\mathbf{C}^{(m)}_w)^{-1} \left[\mathbf{s}^{(m)}_{\boldsymbol{\theta}^{(m)}} - \mathbf{s}^{(m)}_{\boldsymbol{\theta}^{(m)}}\right] = \mathbf{0} \tag{4-13}$$

$$\mathrm{var}\left\{\mathbf{z}^{(m)}\right\} = \mathrm{E}\left\{\mathbf{z}^{(m)} \cdot \mathbf{z}^{(m)T}\right\} = \mathrm{E}\left\{\left[\frac{\partial \ln p(\mathbf{x}^{(m)};\boldsymbol{\theta}^{(m)})}{\partial \boldsymbol{\theta}^{(m)}}\right]\left[\frac{\partial \ln p(\mathbf{x}^{(m)};\boldsymbol{\theta}^{(m)})}{\partial \boldsymbol{\theta}^{(m)}}\right]^T\right\} = \mathbf{J}(\boldsymbol{\theta}^{(m)}) \tag{4-14}$$

Substitution of (4-13) into (4-11) and (4-14) into (4-12) yields

$$\mathrm{E}\left\{\hat{\boldsymbol{\theta}}^{(m)}\right\} = \boldsymbol{\theta}^{(m)} \tag{4-15}$$

and

$$\mathrm{var}\left\{\hat{\boldsymbol{\theta}}^{(m)}\right\} = \mathbf{J}^{-1}(\boldsymbol{\theta}^{(m)}) \tag{4-16}$$

(4-15) and (4-16) indicate that $\boldsymbol{\theta}^{(m)}$ estimated from $\mathbf{r}^{(m)}$ is unbiased and efficient.

The approximate error covariance matrix between different sets of estimators is obtained by

$$\mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(m)}, \hat{\boldsymbol{\theta}}^{(n)}\right\} = \mathbf{J}^{-1}(\boldsymbol{\theta}^{(m)}) \cdot \mathrm{E}\left\{[\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})] \cdot [\mathbf{z}^{(n)}(\boldsymbol{\theta}^{(n)})]^T\right\} \cdot \mathbf{J}^{-1}(\boldsymbol{\theta}^{(n)}) \tag{4-17}$$

Evaluate $\mathrm{E}\left\{[\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})] \cdot [\mathbf{z}^{(n)}(\boldsymbol{\theta}^{(n)})]^T\right\}$ by substituting (4-7) in then

$$\mathrm{E}\left\{[\mathbf{z}^{(m)}(\boldsymbol{\theta}^{(m)})] \cdot [\mathbf{z}^{(n)}(\boldsymbol{\theta}^{(n)})]^T\right\} = \mathrm{E}\left\{\left[\frac{\partial \ln p(\mathbf{x}^{(m)};\boldsymbol{\theta}^{(m)})}{\partial \boldsymbol{\theta}^{(m)}}\right]\left[\frac{\partial \ln p(\mathbf{x}^{(n)};\boldsymbol{\theta}^{(n)})}{\partial \boldsymbol{\theta}^{(n)}}\right]^T\right\} \tag{4-18}$$

(4) DML Estimator and Original Estimator

Let $\boldsymbol{\theta}^{(m)} = \mathbf{A}_m \boldsymbol{\theta}$, then $\mathbf{A}_m$ is the matrix of zero and one entries in the appropriate locations. Let $\boldsymbol{\varphi} = [\boldsymbol{\theta}^{(1)^T}, \boldsymbol{\theta}^{(2)^T}, ..., \boldsymbol{\theta}^{(M)^T}]^T$, then $\boldsymbol{\varphi} = \mathbf{A}\boldsymbol{\theta}$, where $\mathbf{A} = [\mathbf{A}_1^T, \mathbf{A}_2^T, ..., \mathbf{A}_2^T]^T$.

Based on the weighted linear LS methods, the estimates of $\boldsymbol{\theta}$ can be gotten from the measurements $\boldsymbol{\varphi}$ by

$$\hat{\boldsymbol{\theta}} = (\mathbf{A}^T\mathbf{W}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}^{-1}\cdot\hat{\boldsymbol{\varphi}} \tag{4-19}$$

where $\mathbf{W} = \mathrm{cov}\{\hat{\boldsymbol{\varphi}}\}$ is composed from the sub-matrices $\mathrm{var}\left\{\hat{\boldsymbol{\theta}}^{(m)}\right\}$ and $\mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(m)},\hat{\boldsymbol{\theta}}^{(n)}\right\}$ given by (4-16) and (4-17) as

$$\begin{aligned}\mathbf{W} &= \mathrm{cov}\{\hat{\boldsymbol{\varphi}}\} = \mathrm{E}\left\{\hat{\boldsymbol{\varphi}}\hat{\boldsymbol{\varphi}}^T\right\} \\ &= \begin{bmatrix} \mathrm{var}\left\{\hat{\boldsymbol{\theta}}^{(1)}\right\} & \mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(1)},\hat{\boldsymbol{\theta}}^{(2)}\right\} & \cdots & \mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(1)},\hat{\boldsymbol{\theta}}^{(M)}\right\} \\ \mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(2)},\hat{\boldsymbol{\theta}}^{(1)}\right\} & \mathrm{var}\left\{\hat{\boldsymbol{\theta}}^{(2)}\right\} & \cdots & \mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(2)},\hat{\boldsymbol{\theta}}^{(M)}\right\} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(M)},\hat{\boldsymbol{\theta}}^{(1)}\right\} & \mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(M)},\hat{\boldsymbol{\theta}}^{(2)}\right\} & \cdots & \mathrm{var}\left\{\hat{\boldsymbol{\theta}}^{(M)}\right\} \end{bmatrix}\end{aligned} \tag{4-20}$$

Calculate the bias and error covariance matrix of $\hat{\boldsymbol{\theta}}$ to evaluate the efficiency. We already know the components of $\hat{\boldsymbol{\varphi}}$ are efficient from (4-15) and (4-16). Then $\mathrm{E}\{\hat{\boldsymbol{\varphi}}\} = \mathbf{A}\boldsymbol{\theta}$ and

$$\mathrm{E}_{\mathrm{DML}}\left\{\hat{\boldsymbol{\theta}}\right\} = (\mathbf{A}^T\mathbf{W}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}^{-1}\cdot\mathrm{E}\{\hat{\boldsymbol{\varphi}}\} = (\mathbf{A}^T\mathbf{W}^{-1}\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{W}^{-1}\mathbf{A})\boldsymbol{\theta} = \boldsymbol{\theta} \tag{4-21}$$

$$\mathrm{cov}_{\mathrm{DML}}\left\{\hat{\boldsymbol{\theta}}\right\} = (\mathbf{A}^T\mathbf{W}^{-1}\mathbf{A})^{-1} \tag{4-22}$$

The $(i, j)^{th}$ element of FIM of $\boldsymbol{\theta}$ is given by

$$\mathbf{J}_{ij}(\boldsymbol{\theta}) = -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\right\} \tag{4-23}$$

If the $\boldsymbol{\theta}^{(m)}$ is continually taken from $\boldsymbol{\theta}$, which means $\mathbf{A}$ is identical matrix and $\boldsymbol{\varphi} = \boldsymbol{\theta}$. Then

$$\mathrm{cov}_{\mathrm{DML}}\left\{\hat{\boldsymbol{\theta}}\right\} = \mathbf{W} \tag{4-24}$$

If the various $\mathbf{x}^{(m)}$ constitute disjoint of data vector $\mathbf{x}$, $\mathrm{cov}\{\hat{\boldsymbol{\varphi}}\}$ becomes block diagonal and (4-22) has the form

$$\mathrm{cov}_{\mathrm{DML}}\left\{\hat{\boldsymbol{\theta}}\right\} = \left[\sum_{m=1}^{M} \mathbf{A}_m^T \mathbf{J}(\boldsymbol{\theta}^{(m)}) \mathbf{A}_m\right]^{-1} \tag{4-25}$$

And

$$p(\mathbf{x};\boldsymbol{\theta}) = \prod_{m=1}^{M} p(\mathbf{x}^{(m)};\boldsymbol{\theta}^{(m)}) \tag{4-26}$$

and

$$\mathbf{J}_{ij}(\boldsymbol{\theta}) = -\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial\theta_i\,\partial\theta_j}\right\} = -\sum_{m=1}^{M}\mathrm{E}\left\{\frac{\partial^2 \ln p(\mathbf{x}^{(m)};\boldsymbol{\theta}^{(m)})}{\partial\theta_i\,\partial\theta_j}\right\} = \sum_{m=1}^{M}\mathbf{A}_m^T\mathbf{J}(\boldsymbol{\theta}^{(m)})\mathbf{A}_m \tag{4-27}$$

Therefore

$$\mathrm{cov}_{\mathrm{DML}}\{\hat{\boldsymbol{\theta}}\} = \mathbf{J}^{-1}(\boldsymbol{\theta}) \tag{4-28}$$

(4-21) and (4-28) show that if $\mathbf{x}^{(m)}$ constitute disjoint, $\hat{\boldsymbol{\theta}}$ is efficient. Decentralizing the estimation procedure does not affect performance level.

But in general, $\mathbf{x}^{(m)}$ is not always disjoint, or there is correlation between estimate sets. DML maybe introduce degradation in performance.

6.1.2 DML TDOA/FDOA Estimates

Assume there are total *M* sensors, without loss of generality, let M^{th} sensor be the reference one. $\boldsymbol{\theta} = [\tau_{1,M}, \upsilon_{1,M}, ..., \tau_{M-1,M}, \upsilon_{M-1,M}]^T$, $\boldsymbol{\theta}^{(m)} = \boldsymbol{\theta}_m = [\tau_{m,M}, \upsilon_{m,M}]^T, m = 1, ..., M-1$.

Sensor received signal as

$$x_m(t) = s(t;\boldsymbol{\theta}_m) + w_m(t), \quad m = 1, 2, ..., M-1 \tag{4-29}$$

and

$$x_M(t) = s(t;\boldsymbol{\theta}) + w_M(t) \tag{4-30}$$

Let $\mathbf{x}^{(m)} = [\mathbf{x}_m^T, \mathbf{x}_M^T]^T$, then $\mathbf{x}^{(m)}$ is not disjoint, (4-28) maybe not hold. Define

$$q_m(\mathbf{x}_m;\boldsymbol{\theta}_m) = -\frac{1}{2}(\mathbf{x}_m - \mathbf{s}_{\boldsymbol{\theta}_m})^T \mathbf{C}_m^{-1}(\mathbf{x}_m - \mathbf{s}_{\boldsymbol{\theta}_m}) \tag{4-31}$$

Then

$$\ln p(\mathbf{x};\boldsymbol{\theta}) = \sum_{m=1}^{M} q_m(\mathbf{x}_m;\boldsymbol{\theta}_m) \tag{4-32}$$

Here we ignore the $\frac{1}{(2\pi)^{N/2}\det^{1/2}(\mathbf{C}_w)}$ part, since this part is zero when we take the derivative of $p(\mathbf{x};\boldsymbol{\theta})$. According to [13]

$$\mathbf{J}_m = -\mathrm{E}\left\{\begin{bmatrix} \frac{\partial^2 q_m}{\partial \tau_{m,M}{}^2} & \frac{\partial^2 q_m}{\partial \tau_{m,M}\partial v_{m,M}} \\ \frac{\partial^2 q_m}{\partial \tau_{m,M}\partial v_{m,M}} & \frac{\partial^2 q_m}{\partial v_{m,M}{}^2} \end{bmatrix}\right\} = (B\cdot T\cdot \gamma_m)\begin{bmatrix} \beta & 0 \\ 0 & T_e \end{bmatrix} = \gamma_m \mathbf{F} \tag{4-33}$$

where B is noise bandwidth at receiver input, assumed same for all the receivers; β is "rms radian frequency " calculated by

$$\beta = (2\pi)^2 \left[\frac{\int_{-\infty}^{\infty} f^2 |S(f)|^2 \, df}{\int_{-\infty}^{\infty} |S(f)|^2 \, df}\right] \tag{4-34}$$

T_e is "rms integration time" by

$$T_e = (2\pi)^2 \left[\frac{\int_{-\infty}^{\infty} t^2 |s(t)|^2 \, dt}{\int_{-\infty}^{\infty} |s(t)|^2 \, dt}\right] \tag{4-35}$$

γ_m is the input signal noise ratio. And define $\mathbf{F} = BT\begin{bmatrix} \beta & 0 \\ 0 & T_e \end{bmatrix}$, and assume this matrix is the same for all receivers.

And define

$$\mathbf{J}_{m,n} = -\mathrm{E}\left\{\begin{bmatrix} \dfrac{\partial^2 q_M}{\partial\tau_{m,M}\partial\tau_{n,M}} & \dfrac{\partial^2 q_M}{\partial\tau_{m,M}\partial v_{n,M}} \\ \dfrac{\partial^2 q_M}{\partial\upsilon_{m,M}\partial\tau_{n,M}} & \dfrac{\partial^2 q_M}{\partial\upsilon_{m,M}\partial v_{n,M}} \end{bmatrix}\right\} = (B\cdot T\cdot\gamma_M)\begin{bmatrix}\beta & 0\\ 0 & T_e\end{bmatrix} = \gamma_M\mathbf{F},\quad m\neq n \tag{4-36}$$

Then for the centralized ML method,

$$\begin{aligned}&\mathrm{cov}_{\mathrm{CML}}(\hat{\boldsymbol{\theta}}) = \mathbf{J}^{-1}(\boldsymbol{\theta})\\ &= \begin{bmatrix} (\gamma_1+\gamma_M)\mathbf{F} & \gamma_M\mathbf{F} & \cdots & \gamma_M\mathbf{F}\\ \gamma_M\mathbf{F} & (\gamma_2+\gamma_M)\mathbf{F} & \cdots & \gamma_M\mathbf{F}\\ \cdots & \cdots & \ddots & \vdots\\ \gamma_M\mathbf{F} & \gamma_M\mathbf{F} & \cdots & (\gamma_{M-1}+\gamma_M)\mathbf{F}\end{bmatrix}^{-1}_{2(M-1)\times 2(M-1)}\end{aligned} \tag{4-37}$$

From the above discussing, we know for the decentralized ML method the covariance matrix of each ML is

$$\mathrm{var}_{\mathrm{DML}}\left\{\hat{\boldsymbol{\theta}}^{(m)}\right\} = \mathbf{J}^{-1}(\boldsymbol{\theta}^{(m)}) = \left[\mathbf{J}_m + \mathbf{J}_M\right]^{-1} = \frac{1}{(\gamma_1+\gamma_M)}\mathbf{F}^{-1} \tag{4-38}$$

Substitute (4-31) into (4-18), get

$$\mathrm{cov}\left\{\mathbf{z}^{(m)},\mathbf{z}^{(n)}\right\} = -\mathrm{E}\left\{\begin{bmatrix} \dfrac{\partial^2 q_M}{\partial\tau_{m,M}\partial\tau_{n,M}} & \dfrac{\partial^2 q_M}{\partial\tau_{m,M}\partial v_{n,M}} \\ \dfrac{\partial^2 q_M}{\partial\upsilon_{m,M}\partial\tau_{n,M}} & \dfrac{\partial^2 q_M}{\partial\upsilon_{m,M}\partial v_{n,M}} \end{bmatrix}\right\} = \mathbf{J}_{m,n},\quad m\neq n \tag{4-39}$$

Substitute (4-39) into (4-17), get

$$\mathrm{cov}\left\{\hat{\boldsymbol{\theta}}^{(m)},\hat{\boldsymbol{\theta}}^{(n)}\right\} = \left[\mathbf{J}_m + \mathbf{J}_{m,M}\right]^{-1}\mathbf{J}_{m,n}\left[\mathbf{J}_m + \mathbf{J}_{m,M}\right]^{-1} = \frac{\gamma_M}{(\gamma_m+\gamma_M)(\gamma_n+\gamma_M)}\mathbf{F}^{-1} \tag{4-40}$$

Since the $\boldsymbol{\theta}^{(m)}$ is continually taken from $\boldsymbol{\theta}$ and disjointed, then

$$\mathrm{cov}_{\mathrm{DML}}(\hat{\boldsymbol{\theta}}) = \begin{bmatrix} \dfrac{1}{(\gamma_1+\gamma_M)}\mathbf{F}^{-1} & \cdots & \dfrac{\gamma_M}{(\gamma_1+\gamma_M)(\gamma_{M-1}+\gamma_M)}\mathbf{F}^{-1}\\ \vdots & \ddots & \vdots\\ \dfrac{\gamma_M}{(\gamma_{M-1}+\gamma_M)(\gamma_1+\gamma_M)}\mathbf{F}^{-1} & \cdots & \dfrac{1}{(\gamma_{M-1}+\gamma_M)}\mathbf{F}^{-1}\end{bmatrix}_{2(M-1)\times 2(M-1)} \tag{4-41}$$

If assume all the $\gamma_m \equiv \gamma$, then (4-41) and (4-37) becomes to

$$\mathrm{var}_{\mathrm{DML}}(\hat{\boldsymbol{\theta}}) = \frac{1}{4\gamma}\begin{bmatrix} 2\mathbf{F}^{-1} & \cdots & \mathbf{F}^{-1} \\ \vdots & \ddots & \vdots \\ \mathbf{F}^{-1} & \cdots & 2\mathbf{F}^{-1} \end{bmatrix}_{2(M-1)\times 2(M-1)} \tag{4-42}$$

and based on matrix inverse calculation

$$\mathrm{var}_{\mathrm{CML}}(\hat{\boldsymbol{\theta}}) = \frac{1}{4\gamma}\frac{M-1}{M}\begin{bmatrix} 2\mathbf{F}^{-1} & \cdots & \mathbf{F}^{-1} \\ \vdots & \ddots & \vdots \\ \mathbf{F}^{-1} & \cdots & 2\mathbf{F}^{-1} \end{bmatrix}_{2(M-1)\times 2(M-1)} \tag{4-43}$$

Compare (4-43) and (4-42), which are gotten by centralized and de-centralized methods respectively, we can see centralized method has smaller variances than de-centralized one (since $(M-1/M)<1$). That is the possible improvements in the estimate of $\boldsymbol{\theta}_m$ from use of receivers outside the m^{th} pair. But for large number of *M*, this improvement is negligible compare with the computation complexity of high-dimensional cross-correlation. So practically, DML is widely used in the first stage TDOA/FDOA estimation.

6.2 Next Optimal State

We know the estimation accuracy of emitter location is specified by the FIM of $\mathbf{p}_e$ as $\mathbf{J}_{geo}(\mathbf{p}_e) = \mathbf{H}^T\mathbf{F}(\boldsymbol{\theta})\mathbf{H}$. The matrices are defined as before. To get optimal $\mathbf{J}_{geo}(\mathbf{p}_e)$, we can consider $\mathbf{F}(\boldsymbol{\theta})$ and $\mathbf{H}$ separately, since they are related with different parameters. $\mathbf{F}(\boldsymbol{\theta})$ is only depend on the signal model and received signal data quality, $\mathbf{F}(\boldsymbol{\theta})$ can be optimally compressed [23]. $\mathbf{H}$ is totally specified by the current sensor local states and referenced emitter location. So it is possible to find the next step for sensors based on current data quality and relationship of the paired sensors. For simplicity, we talk about one pair first.

For one pair

$$\mathbf{J}_m = \mathbf{G}_m^T \mathbf{F}_m \mathbf{G}_m = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}$$
$$J_{11} = f_t g_{11}^2 + f_v g_{21}^2 + 2 f_{\tau,\upsilon} g_{11} g_{21} \tag{4-44}$$
$$J_{22} = f_t g_{12}^2 + f_v g_{22}^2 + 2 f_{\tau,\upsilon} g_{12} g_{22}$$
$$J_{12} = J_{21} = f_t g_{11} g_{12} + f_v g_{21} g_{22} + f_{\tau,\upsilon} g_{11} g_{22}$$

where

$$\mathbf{G}_m = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \quad \mathbf{F}_m = \begin{bmatrix} f_\tau & f_{\tau,\upsilon} \\ f_{\tau,\upsilon} & f_\upsilon \end{bmatrix}$$
$$g_{11} = \frac{1}{c}\left(\frac{x_1 - x_e}{r_1} - \frac{x_2 - x_e}{r_2} \right)$$
$$g_{12} = \frac{1}{c}\left(\frac{y_1 - y_e}{r_1} - \frac{y_2 - y_e}{r_2} \right) \tag{4-45}$$
$$g_{21} = \frac{fe}{c}\left(\left[\frac{(x_1 - x_e)(\mathbf{u}_1^T \dot{\mathbf{s}}_1)}{r_1^2} - \frac{\dot{x}_1}{r_1} \right] - \left[\frac{(x_2 - x_e)(\mathbf{u}_2^T \dot{\mathbf{s}}_2)}{r_2^2} - \frac{\dot{x}_2}{r_2} \right] \right)$$
$$g_{22} = \frac{fe}{c}\left(\left[\frac{(y_1 - y_e)(\mathbf{u}_1^T \dot{\mathbf{s}}_1)}{r_1^2} - \frac{\dot{y}_1}{r_1} \right] - \left[\frac{(y_2 - y_e)(\mathbf{u}_2^T \dot{\mathbf{s}}_2)}{r_2^2} - \frac{\dot{y}_2}{r_2} \right] \right)$$

In chapter 4, the selection object is as

$$\boldsymbol{\beta}^* = \arg \min_{\text{all the given } \boldsymbol{\beta}} \left\{ \text{some scalar value function of CRLB}_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}; \boldsymbol{\beta}) \right\} \tag{4-46}$$

For the selection problem, argument $\boldsymbol{\beta}$ is given and our task is to use some optimal method to find the best one among all the given $\boldsymbol{\beta}$. For the next optimal state problem, argument $\boldsymbol{\beta}$ is specified by some researchable set, which is a set of sensor's future position and velocity and it is specified by the sensor's current position, velocity, and acceleration and time limitation. So rewrite (4-46) as

$$\boldsymbol{\beta}^* = \arg \min_{\text{researchable set of } \boldsymbol{\beta}} \left\{ \text{some scalar value function of CRLB}_{\boldsymbol{\alpha}}(\boldsymbol{\alpha}; \boldsymbol{\beta}) \right\} \tag{4-47}$$

And as discussed in chapter 4, the CRLB of emitter location defines an error ellipsoid. Our final goal is to minimize some measurement of the error ellipsoid, for example, the minimization can be towards the determinant or the trace of the CRLB. That is to define a scalar function f of CRLB or FIM as $f_{\mathbf{J}}(\mathbf{J}_{geo})$. $f_{\mathbf{J}}(\mathbf{J}_{geo})$ is also a scalar function of position vector $\mathbf{s}$ and velocity vector $\dot{\mathbf{s}}$ as $f_{\mathbf{J}}(\mathbf{s},\dot{\mathbf{s}})$. Rewrite (4-47) as

$$(\mathbf{s}^*, \dot{\mathbf{s}}^*) = \arg \min_{\substack{\text{reachable set} \\ \text{of } \mathbf{s} \text{ and } \dot{\mathbf{s}}}} \left\{ f_{\mathbf{J}}(\mathbf{s},\dot{\mathbf{s}}) \right\} \tag{4-48}$$

6.2.1 Optimal Criterion

The relationship of trace and determinant of CRLB and FIM is as following

$$\begin{aligned} \text{trace}(\mathbf{CRLB}_{geo}) &= \frac{\text{trace}(\mathbf{J}_{geo})}{\det(\mathbf{J}_{geo})} \\ \det(\mathbf{CRLB}_{geo}) &= \frac{1}{\det(\mathbf{J}_{geo})} \end{aligned} \tag{4-49}$$

For two pairs FIM of $\mathbf{p}_e$ is the sum of two items assuming there is no sensor sharing

$$\mathbf{J}_{geo} = \mathbf{J}_1 + \mathbf{J}_2 \tag{4-50}$$

Then

$$\text{trace}(\mathbf{J}_{geo}) = \text{trace}(\mathbf{J}_1) + \text{trace}(\mathbf{J}_2) \tag{4-51}$$

And from Minkowski inequality

$$\det(\mathbf{J}_{geo}) \geq \det(\mathbf{J}_1) + \det(\mathbf{J}_2) \tag{4-52}$$

Also

$$\mathbf{J}_{geo}^{-1} = \mathbf{J}_1^{-1} - \mathbf{J}_1^{-1}(\mathbf{J}_1^{-1} + \mathbf{J}_2^{-1})^{-1}\mathbf{J}_1^{-1} \tag{4-53}$$

The expression of objective as a function of position and velocity will be too complicated to get an explicit solution. We can always use the numerical method to get

the optimal one, but it is time consuming. So following we only discuss about some sub-optimal solutions which may have a clear solution expression. Since the cross term $f_{\tau,\upsilon}$ is negligible compare with others, so we ignore this one at the following discussion.

6.2.2 Minimize Trace of $\mathbf{J}_{geo}^{-1}$

As we discussed before, maximizing the trace of FIM tends to minimize the trace of the CRLB. Objective function rewritten as

$$\max_{\text{reachable set of } \mathbf{s} \text{ and } \dot{\mathbf{s}}} \left\{ f_{\mathbf{J}}(\mathbf{s},\dot{\mathbf{s}}) = \text{trace}(\mathbf{J}_{geo}) \right\} \tag{4-54}$$

and

$$f_{\mathbf{J}}(\mathbf{s},\dot{\mathbf{s}}) = \text{trace}(\mathbf{J}_{geo}) = k_\tau \cdot g_\tau + k_\upsilon \cdot g_\upsilon \tag{4-55}$$

where

$$\begin{aligned} & k_\tau = f_\tau \frac{2}{c^2} \quad k_\upsilon = f_\upsilon (\frac{f_e}{c})^2 \\ & g_\tau = 1 - \mathbf{u}_1^T \mathbf{u}_2 \\ & g_\upsilon = \frac{1}{r_1^2} \dot{\mathbf{s}}_1^T (\mathbf{I} - \mathbf{u}_1 \mathbf{u}_1^T) \dot{\mathbf{s}}_1 + \frac{1}{r_2^2} \dot{\mathbf{s}}_2^T (\mathbf{I} - \mathbf{u}_2 \mathbf{u}_2^T) \dot{\mathbf{s}}_2 \\ & \qquad + \frac{2}{r_1 r_2} \dot{\mathbf{s}}_1^T (\mathbf{u}_1 \mathbf{u}_1^T + \mathbf{u}_2 \mathbf{u}_2^T - \mathbf{u}_1 \mathbf{u}_1^T \mathbf{u}_2 \mathbf{u}_2^T - \mathbf{I}) \dot{\mathbf{s}}_2 \end{aligned} \tag{4-56}$$

We can see g_υ is a complex combination of $\mathbf{s}$ and $\dot{\mathbf{s}}$.

(1) Optimal with respect to velocity only

Rewrite (4-55) as a function of $\dot{\mathbf{s}}$ only

$$f_{\mathbf{J}}(\dot{\mathbf{s}}) = \dot{\mathbf{s}}_1^T \mathbf{A}_1 \dot{\mathbf{s}}_1 + \dot{\mathbf{s}}_2^T \mathbf{A}_2 \dot{\mathbf{s}}_2 + 2\dot{\mathbf{s}}_1^T \mathbf{B} \dot{\mathbf{s}}_2 + k_\tau \tag{4-57}$$

where

$$\mathbf{A}_1 = \frac{k_v}{r_1^2}(\mathbf{I} - \mathbf{u}_1\mathbf{u}_1^T)$$
$$\mathbf{A}_2 = \frac{k_v}{r_2^2}(\mathbf{I} - \mathbf{u}_2\mathbf{u}_2^T) \tag{4-58}$$
$$\mathbf{B} = \frac{k_v}{r_1 r_2}[\,\mathbf{u}_1\mathbf{u}_1^T + \mathbf{u}_2\mathbf{u}_2^T - \mathbf{u}_1\mathbf{u}_1^T\mathbf{u}_2\mathbf{u}_2^T - \mathbf{I}\,]$$

It is clear that $f_{\mathbf{J}}(\dot{\mathbf{s}})$ is a quadratic form of $\dot{\mathbf{s}}$. $\mathbf{A}_1$ and $\mathbf{A}_2$ are positive semi-definite matrices. So it is a convex problem. Since the maximal value of a convex problem is on the boundary of the feasible set, if the feasible set is also convex. So the solution for this sub-optimal problem is on the boundary of the feasible velocity set.

(2) Optimal with respect to position only

Since the next optimal state is within a few seconds, the small change of r_i can be ignored for computation simplicity. Rewrite (4-55) as a function of $\mathbf{s}$ only.

$$f_{\mathbf{J}}(\mathbf{s}) = \mathbf{s}_1^T\mathbf{A}_1\mathbf{s}_1 + \mathbf{s}_2^T\mathbf{A}_2\mathbf{s}_2 - \frac{k_\tau}{r_1 r_2}\mathbf{s}_1^T\mathbf{s}_2 - \frac{2k_\upsilon}{r_1^2 r_2^2}\dot{\mathbf{s}}_1^T\mathbf{s}_1\mathbf{s}_1^T\mathbf{s}_2\mathbf{s}_2^T\dot{\mathbf{s}}_2 + k \tag{4-59}$$

where

$$k = 1 + k_v\left(\frac{1}{r_1^2}\dot{\mathbf{s}}_1^T\dot{\mathbf{s}}_1 + \frac{1}{r_2^2}\dot{\mathbf{s}}_2^T\dot{\mathbf{s}}_2 + \frac{2}{r_1 r_2}\dot{\mathbf{s}}_1^T\dot{\mathbf{s}}_2\right)$$
$$\mathbf{A}_1 = k_v\left(\frac{2}{r_1^3 r_2}\dot{\mathbf{s}}_1\dot{\mathbf{s}}_2^T - \frac{1}{r_1^4}\dot{\mathbf{s}}_1\dot{\mathbf{s}}_1^T\right) \tag{4-60}$$
$$\mathbf{A}_2 = k_v\left(\frac{2}{r_1 r_2^3}\dot{\mathbf{s}}_2\dot{\mathbf{s}}_1^T - \frac{1}{r_2^4}\dot{\mathbf{s}}_2\dot{\mathbf{s}}_2^T\right)$$

The properties of $\mathbf{A}_1$ and $\mathbf{A}_2$ depend on the values of distances and velocities. And also the convexity of this function is not obvious. There is not an explicit solution for the objective function. We have to use the numerical method to find out the optimal value of the objective function.

6.2.3 Minimize Determinant of $\mathbf{J}_{geo}^{-1}$

Since $\det(\mathbf{J}_{geo}^{-1}) = 1/\det(\mathbf{J}_{geo})$, then the minimization of determinant of $\mathbf{J}_{geo}^{-1}$ is equal to the maximization of determinant of $\mathbf{J}_{geo}$. Objective function rewritten as

$$(\mathbf{s}^*, \dot{\mathbf{s}}^*) = \arg \max_{\text{reachable set of } \mathbf{s} \text{ and } \dot{\mathbf{s}}} \left\{ f_{\mathbf{J}}(\mathbf{s}, \dot{\mathbf{s}}) = \det(\mathbf{J}_{geo}) \right\} \tag{4-61}$$

And

$$\begin{aligned} f_{\mathbf{J}}(\mathbf{s}, \dot{\mathbf{s}}) &= \det(\mathbf{J}_{geo}) \\ &= f_t f_v (g_{11} g_{22} - g_{12} g_{21})^2 \\ &= \det(\mathbf{F}) \cdot \det^2(\mathbf{G}(\mathbf{s}, \dot{\mathbf{s}})) \end{aligned} \tag{4-62}$$

$\mathbf{G}(\mathbf{s}, \dot{\mathbf{s}})$ means matrix $\mathbf{G}$ is a function of $(\mathbf{s}, \dot{\mathbf{s}})$. Since $\det(\mathbf{G})$ can be positive or negative, rewrite (4-61) as

$$(\mathbf{s}^*, \dot{\mathbf{s}}^*) = \arg \max_{\text{reachable set of } \mathbf{s} \text{ and } \dot{\mathbf{s}}} \left\{ (-1)^{\text{sign}(|\mathbf{G}(\mathbf{s}, \dot{\mathbf{s}})|)} \cdot |\mathbf{G}(\mathbf{s}, \dot{\mathbf{s}})| \right\} \tag{4-63}$$

where

$$\text{sign}\left(|\mathbf{G}(\mathbf{s}, \dot{\mathbf{s}})|\right) = \begin{cases} 0, \text{ if } |\mathbf{G}(\mathbf{s}, \dot{\mathbf{s}})| > 0 \\ 1, \text{ if } |\mathbf{G}(\mathbf{s}, \dot{\mathbf{s}})| < 0 \end{cases} \tag{4-64}$$

(1) Optimal with respect to velocity only

Write the elements of $\mathbf{G}$ with respect to velocity as

$$\begin{aligned} g_{21} &= \left[\mathbf{a}_x^T, -\mathbf{b}_x^T \right] \begin{bmatrix} \dot{\mathbf{s}}_1 \\ \dot{\mathbf{s}}_2 \end{bmatrix} \\ g_{22} &= \left[\mathbf{a}_y^T, -\mathbf{b}_y^T \right] \begin{bmatrix} \dot{\mathbf{s}}_1 \\ \dot{\mathbf{s}}_2 \end{bmatrix} \end{aligned} \tag{4-65}$$

where

$$\mathbf{a}_x = \begin{bmatrix} \frac{r_{1x}^2}{r_1^2} - \frac{1}{r_1} \\ \frac{r_{1x} r_{1y}}{r_1^2} \end{bmatrix} \quad \mathbf{a}_y = \begin{bmatrix} \frac{r_{1x} r_{1y}}{r_1^2} \\ \frac{r_{1y}^2}{r_1^2} - \frac{1}{r_1} \end{bmatrix} \quad \mathbf{b}_x = \begin{bmatrix} \frac{r_{2x}^2}{r_2^2} - \frac{1}{r_2} \\ \frac{r_{2x} r_{2y}}{r_2^2} \end{bmatrix} \quad \mathbf{b}_y = \begin{bmatrix} \frac{r_{2x} r_{2y}}{r_2^2} \\ \frac{r_{2y}^2}{r_2^2} - \frac{1}{r_2} \end{bmatrix} \tag{4-66}$$

Then $|\mathbf{G}|$ as a function of velocity can be written as

$$|\mathbf{G}(\dot{\mathbf{s}})| = \boldsymbol{\beta} \begin{bmatrix} \dot{\mathbf{s}}_1 \\ \dot{\mathbf{s}}_2 \end{bmatrix} \tag{4-67}$$

where

$$\boldsymbol{\beta} = \left[(g_{11}\mathbf{a}_y - g_{12}\mathbf{a}_x)^T, (g_{11}\mathbf{b}_y - g_{12}\mathbf{b}_x)^T \right] \tag{4-68}$$

Therefore $|\mathbf{G}|$ is a linear function of velocity. To find the maximal or minimal value of $|\mathbf{G}|$, we can just evaluate the boundary values of the reachable set. And find the one satisfies (4-63).

(2) Optimal with respect to position only

Also ignore the small change of r_i for computation simplicity. Rewrite $|\mathbf{G}|$ as a function of $\mathbf{s}$ only is a much more complicated than (4-59), There is not an explicit solution for the objective function. We have to use the numerical method to find out the optimal value of the objective function.

6.3 Sensor Error Effects on Next Optimal State Solution

Assuming knowledge of $\mathbf{p}_e$, the influence of any objective function $f_{\mathbf{J}}(\mathbf{s}, \dot{\mathbf{s}})$ by $\mathbf{s}$ and $\dot{\mathbf{s}}$ can be seen from

$$df_{\mathbf{J}} = \left(\frac{\partial f_{\mathbf{J}}}{\partial \mathbf{s}}\right)^T d\mathbf{s} + \left(\frac{\partial f_{\mathbf{J}}}{\partial \dot{\mathbf{s}}}\right)^T d\dot{\mathbf{s}} \tag{3-69}$$

Since r_i is in the order of km, for simplicity, assume small error of $\mathbf{s}$ will not effect r_i.

6.3.1 Uncertainty on Trace of $\mathbf{J}_{geo}$

Rewrite (3-69) as

$$df_{\mathbf{J}} = k_\tau \left(\frac{\partial g_\tau}{\partial \mathbf{s}}\right)^T d\mathbf{s} + k_\upsilon \left(\frac{\partial g_\upsilon}{\partial \mathbf{s}}\right)^T d\mathbf{s} + k_\upsilon \left(\frac{\partial g_\upsilon}{\partial \dot{\mathbf{s}}}\right)^T d\dot{\mathbf{s}} \tag{4-70}$$

where k_τ, k_υ, g_τ and g_υ are defined in (4-56).

$$\frac{\partial g_\tau}{\partial \mathbf{s}} = \frac{-2}{r_1 r_2}\begin{bmatrix}\mathbf{r}_2 \\ \mathbf{r}_1\end{bmatrix} \tag{4-71}$$

The change in $\mathbf{s}$ is reflected in g_τ through

$$\left\|\frac{\partial g_\tau}{\partial \mathbf{s}}\right\| = \frac{2}{r_1 r_2}\left\|\begin{matrix}\mathbf{r}_2 \\ \mathbf{r}_1\end{matrix}\right\| \tag{4-72}$$

Therefore $\| \partial g_\tau / \partial \mathbf{s} \|$ is on the order of $1/r_i$.

And

$$\frac{\partial g_\upsilon}{\partial \mathbf{s}} = \begin{bmatrix} \dfrac{2(\dot{\mathbf{s}}_1\dot{\mathbf{s}}_1^T)\mathbf{r}_1}{r_1^4} - \dfrac{2(\dot{\mathbf{s}}_1\dot{\mathbf{s}}_2^T)\mathbf{r}_1}{r_1^3 r_2} + \dfrac{2(\dot{\mathbf{s}}_1+\dot{\mathbf{s}}_2)\mathbf{r}_2^T(\dot{\mathbf{s}}_1+\dot{\mathbf{s}}_2)}{r_1^2 r_2^2} \\ \dfrac{2(\dot{\mathbf{s}}_2\dot{\mathbf{s}}_2^T)\mathbf{r}_2}{r_2^4} - \dfrac{2(\dot{\mathbf{s}}_1\dot{\mathbf{s}}_2^T)\mathbf{r}_2}{r_1 r_2^3} + \dfrac{2(\dot{\mathbf{s}}_1+\dot{\mathbf{s}}_2)\mathbf{r}_1^T(\dot{\mathbf{s}}_1+\dot{\mathbf{s}}_2)}{r_1^2 r_2^2} \end{bmatrix} \tag{4-73}$$

Therefore $\| \partial g_\upsilon / \partial \mathbf{s} \|$ is on the order of $1/r_i^3$.

And

$$\frac{\partial g_\upsilon}{\partial \dot{\mathbf{s}}} = 2\begin{bmatrix}\mathbf{A}_1 & \mathbf{B} \\ \mathbf{B}^T & \mathbf{A}_2\end{bmatrix}\begin{bmatrix}\dot{\mathbf{s}}_1 \\ \dot{\mathbf{s}}_2\end{bmatrix} \tag{4-74}$$

$\| \partial g_\upsilon / \partial \dot{\mathbf{s}} \|$ is on the order of $(\| \dot{\mathbf{s}}_i \| / r_i^2)$. Substitute all the order into (4-70), we get

$$\|df_{\mathbf{J}}\| \approx f_\tau \frac{2}{c^2} \cdot \frac{\|d\mathbf{s}\|}{r_k} + f_\upsilon (\frac{f_e}{c})^2 \cdot \frac{\|d\mathbf{s}\|}{r_k^3} + f_\upsilon (\frac{f_e}{c})^2 \cdot \frac{\|d\dot{\mathbf{s}}\|}{r_k^2} \tag{4-75}$$

And we know $\|d\mathbf{s}\|$ is around 100*meters*, and $\|d\dot{\mathbf{s}}\|$ is about 0.02 *meters/second.* Based on all these units/magnitude, we can get an approximate conclusion, the uncertainty of $\dot{\mathbf{s}}$ is more affective than the uncertainty of $\mathbf{s}$ in the computation of trace of FIM. So following we only discuss about the effect of velocity uncertainty on 'next optimal state' solution.

Due the sensor velocity uncertainty, for the objective function as in (4-54), in fact, we are solving

$$\max_{\text{reachable set of } \mathbf{s} \text{ and } \dot{\mathbf{s}}} \{ k_\tau \cdot g_\tau(\mathbf{s}) + k_\upsilon \cdot \mathrm{g}_\upsilon(\mathbf{s}, \dot{\mathbf{s}} + \Delta\dot{\mathbf{s}}) \} \tag{4-76}$$

where

$$\begin{aligned} &\mathrm{g}_\upsilon(\dot{\mathbf{s}} + \Delta\dot{\mathbf{s}}) = \mathrm{g}_\upsilon(\dot{\mathbf{s}}) + \Delta g(\dot{\mathbf{s}}, \Delta\dot{\mathbf{s}}) \\ &\Delta g(\dot{\mathbf{s}}, \Delta\dot{\mathbf{s}}) = k[2\dot{\mathbf{s}}_1^T(\mathbf{A}_1\Delta\dot{\mathbf{s}}_1 + \mathbf{B}\Delta\dot{\mathbf{s}}_2) + 2\dot{\mathbf{s}}_2^T(\mathbf{A}_2\Delta\dot{\mathbf{s}}_2 + \mathbf{B}\Delta\dot{\mathbf{s}}_1) \\ &\qquad + (\Delta\dot{\mathbf{s}}_1^T\mathbf{A}_1\Delta\dot{\mathbf{s}}_1 + \Delta\dot{\mathbf{s}}_2^T\mathbf{A}_2\Delta\dot{\mathbf{s}}_2 + 2\Delta\dot{\mathbf{s}}_1^T\mathbf{B}\Delta\dot{\mathbf{s}}_2)] \end{aligned} \tag{4-77}$$

Then

$$\frac{\partial\Delta g}{\partial\dot{\mathbf{s}}} = 2\begin{bmatrix} \mathbf{A}_1 & \mathbf{B} \\ \mathbf{B}^T & \mathbf{A}_2 \end{bmatrix}\begin{bmatrix} \Delta\dot{\mathbf{s}}_1 \\ \Delta\dot{\mathbf{s}}_2 \end{bmatrix} \tag{4-78}$$

From (4-77) $\Delta g(\dot{\mathbf{s}}, \Delta\dot{\mathbf{s}})$ is a linear function of $\dot{\mathbf{s}}$, and $g_\upsilon(\dot{\mathbf{s}}, \Delta\dot{\mathbf{s}})$ is a convex function of $\dot{\mathbf{s}}$. So $\max_{\dot{\mathbf{s}}}\{g_\upsilon(\dot{\mathbf{s}} + \Delta\dot{\mathbf{s}})\}$ will have the same solution as $\max_{\dot{\mathbf{s}}}\{g_\upsilon(\dot{\mathbf{s}})\}$. Which means the uncertainty of $\dot{\mathbf{s}}$ will not effect the sub-optimal solution of next optimal solution of $\dot{\mathbf{s}}$.

6.3.2 Uncertainty on Determinant of $\mathbf{J}_{geo}$

Following the same discuss as the trace of $\mathbf{J}_{geo}$, we can get the same conclusion that for this sub-optimal objective method, the uncertainty of velocity also has no effect on next optimal state solutions.

7 Conclusion and Future Work

This dissertation develops network-wide optimization over a large number of simultaneously participating sensors, enables the sensors to cooperate efficiently and effectively, and exploiting the accuracy for emitter location system. The contributions of this dissertation are : (1) Explore the importance of signal model in the parameter estimation problem; (2) Investigation is performed to find out all the related aspects about the emitter location accuracy, giving out formulas about the relative measurements; (3) From network management point of view, developing various methods to select and pair a subset of sensors to fulfill the system requirement; (4) Modifying the current widely used estimation strategy to mitigate the reduce of geo-location accuracy due to the sensors' navigation uncertainty; (5) Finally, giving the concept of next optimal state of sensors, giving some sub-optimal solutions and discussing the sensor local states uncertainty's effects on the next optimal solutions.

Besides these achievements there are still many details to be considered in the future research to get more optimal solutions.

We are giving out the overall point of view about the emitter location estimation accuracy formulas and explore some explicit relationship between the accuracy and some system parameters. The future work maybe focus on how all these relationship works and try to find out if these is a simpler formula about all these aspects.

The optimal selection and pairing criterion we used in this dissertation is actually a sub-optimal selection. But our main contribution is to give an idea and procedure about how to make the selection and pairing and what we need to consider when doing the

network management. And the simulation result shows even under our sub-optimal solution, there are still many improvements against traditional procedure.

We proposed an updated method to mitigate the reduced accuracy due to the sensors' local states uncertainty. The updated method needs some knowledge of the sensors uncertainty to construct the new weighted matrix. And we also found out the regular total least square, which is widely used in the problem with inaccurate system information, is not suitable for our problem. Future work can focus on how to improve the normal TLS method to fix in our problem, and then to improve the estimation accuracy even without the knowledge of the probability of sensors' error.

For the next optimal state problem, we investigate that the current information of all the measurements can really give some hints about what will be the best state for the next time period. And we simplify the objective function at the final stage of solving the problem. Since we still need the traditional numerical method to find the solution, we can try to find some other better methods, such as the partial swam optimization, which is a popular optimization issue in these days, to fasten the searching step and maybe get a better solution.

Additional, in this dissertation, we consider only about the stationary emitter problem for illustration simplicity. We can expend all of our ideas and algorithms to moving emitter easily. And also we assume the emitter location as deterministic; if it is random we can modify our proposed methods to other estimation methods, such as Bayesian theorem, Kalman Fitter.

Overall, this dissertation explored the network-wide optimal concept about the emitter location problem. And all the ideas can be easily modified to satisfy the practical requirements.

8 References

[1] S. Kay, *Fundamentals of Statistical Signal Processing Volume I: Estimation Theory,* Englewood Cliffs, NJ: Prentice Hall, 1993.

[2] S. Kay, *Fundamentals of Statistical Signal Processing Volume II: Detection Theory,* Englewood Cliffs, NJ: Prentice Hall, 1993.

[3] T. K. Moom and W. C. Stirling, *Mathematical Methods and Algorithms for Signal Processing*, NJ: Prentice Hall, 2000.

[4] B. Poart, *A Course in Digital Signal Processing*, John Wiley & Sons, INC, 1997

[5] John Wiley and Sons, *Combinatorial Optimization*.

[6] W. R. Hann and S. A. Tretter, "Optimum Processing for Delay-Vector Estimation in Passive Signal Arrays," *IEEE Trans. Information Theory*, vol. IT-19, pp. 608 - 614, Sept. 1973.

[7] W. R. Hann, "Optimum Signal Processing for Passive Sonar Range and Bearing Estimation," *J. Acoust. Soc. Am.*, Vol. 58 pp. 201 – 207, July 1975.

[8] Harry B. Lee, "A Novel Procedure for Assessing the Accuracy of Hyperbolic Multilateration Systems," *IEEE Transactions on Aerospace and Electronic Systems*, Vol.AES-11, No.1, January 1975.

[9] C. H. Knapp and G. C. Carter, "The Generalized Correlation Method for Estimation of Time Delay," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-24, pp. 320 - 327, Aug. 1976.

[10] P. M. Schultheis and E. Weinstein, "Estimation of Differential Doppler Shifts," *J. Acoust. Soc. Am.*, Vol. 66 pp. 1412 – 1419, Nov. 1979.

[11] G. C. Carter, "Time Delay Estimation for Passive Sonar Signal Processing," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-29, pp. 463 - 470, June 1981.

[12] A. H. Quazi, "An Overview on the Time Delay Estimate in Active and Passive Systems for Target Localization," *IEEE Trans. Acoust., Speech, and Signal Processing*, vol. ASSP-29, pp. 527 - 533, June 1981.

[13] Seymour Stein, "Algorithms for Ambiguity Function Processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-29, No. 3, June 1981.

[14] Ehud Weistein, "Decentralization of the Gaussian Maximum Likelihood Estimator and Its Applications to Passive Array Processing," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol. ASSP-29, No.5, October 1981.

[15] M. Wax, "The Joint Estimation of Differential Delay, Doppler, and Phase," *IEEE Trans. Information Theory*, vol. IT-28, pp. 817 - 820, Sept. 1982.

[16] Paul C. Chestnut, "Emitter Location Accuracy Using TDOA and Differential Doppler," *IEEE Transactions on Aerospace and Electronic Systems*, VOL.AES-18, No.2, March 1982.

[17] Don J. Torrieri, "Statistical Theory of Passive Location System," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 2, March 1984, pp. 183 – 198.

[18] B. Friedlander, "On the Cramer-Rao Bound for Time Delay and Doppler Estimation," *IEEE Trans. Information Theory*, vol. IT-30, pp. 575 - 580, May 1984.

[19] T. Berger and R. Blahut, "Coherent Estimation of Differential Delay and Differential Doppler," *Proc. ofConf. on Information Sciences and Systems*, Princeton University, pp. 537 – 541,1984.

[20] S. Stein, "Differential delay/Doppler ML Estimation with Unknown Signals," *IEEE Transactions on Signal Processing*, vol. 41, pp. 2717-2719, August 1993.

[21] K. Madsen, H.B. Nielsen, and O. Tingleff, "Methods for Non-Linear Least Squares Problems," 2nd Edition, April 2004

[22] Van Huffle, S., and J. Vandewalle, "On the Accuracy of Total Least Squares and Least Squares Techniques in the Presence of Errors on all Data," Automatic, Vol. 25, No. 5, 1989, pp. 765 – 769.

[23] Mo Chen, "Data Compression For Inference Tasks In Wireless Sensor Networks," Dissertation, 2005, State University of New York at Binghamton

[24] N. Eva Wu and Mark L. Fowler, "An Error Bound for Sensor Fusion with Application to Doppler Frequency Based Emitter Location," *IEEE Transactions on Automatic Control*, vol. 51, no. 4, April 2006, pp. 631 – 635.

[25] Mark L. Fowler and M. Chen, "Fisher-Information-Based Data Compression for Estimation Using Two Sensors," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 3, July 2005, pp. 1131 - 1137.

[26] N. Eva Wu and Mark L. Fowler, "Aperture Error Mitigation via Local-State Estimation for Frequency-Based Emitter Location," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 2, April 2003, pp. 414 – 429.

[27] Mark L. Fowler, "Analysis of Passive Emitter Location using Terrain Data," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 37, no. 2, pp. 495 – 507, April 2001.

[28] Mark L. Fowler and X. Hu, "Signal Models for TDOA/FDOA Estimation," to appear in *IEEE Transactions on Aerospace and Electronic Systems*.

[29] Xi Hu and Mark L. Fowler, "Sensor Selection for Multiple Sensor Emitter Location Systems," *IEEE Aerospace Conference*, Big Sky, Montana, 1 – 8 March, 2008.

[30] Mark L. Fowler and Xi Hu, "The Impact of Acoustic Signal Models on TDOA/FDOA Estimation," *(Invited Paper), Acoustical Society of America*, 154th Meeting, New Orleans, Louisiana, 27 November--1 December 2007.

[31] Xi Hu, Mo Chen, and Mark L. Fowler, "Exploiting Data Compression Methods for Network-Level Management of Multi-Sensor Systems," *Proceedings of SPIE - The International Society for Optical Engineering*, v 6315, Mathematics of Data/Image Pattern Recognition, Compression, and Encryption with Applications IX, 2006.

[32] M. L. Fowler and Mo Chen, "Perspectives on Data Compression for Estimations From Sensors," *Proceedings of SPIE*, 2006.

[33] Mo Chen and Mark L. Fowler, "Data Compression for Simultaneous/Sequential Inference Tasks in Sensor Networks," *(Invited Paper), IEEE International Conference on Acoustics, Speech, and Signal Processing* (*ICASSP*), 2006, Toulouse, France, May 14 – 19, 2006, vol. V, pp. 1057 – 1060.

[34] Mark L. Fowler and M. Chen, "Evaluating Fisher Information From Data for Task-Driven Data Compression," *Proceedings of Conference on Information Sciences and Systems*, Princeton University, March 22-24, 2004, pp. 976 – 972

[35] M. Chen and Mark L. Fowler, "Geometry-Adaptive Data Compression For TDOA/FDOA Location," *IEEE ICASS*P, 2005, Philadelphia, PA, March 18 – 23, 2005, vol. 4, pp. 1069 – 1072.

[36] K. Hints and E. McVey, "A Measure of the Information Gain Attributable to Cueing," *IEEE T. Sys. Man Cyb.*, 1991.

[37] J. Liu, J. Reich, and F. Zhao, "Collaborative In-Network Processing for Target Tracking," *EURASIP JASP: Special Issues on Sensor Networks*, March 2003.

[38] K.C. Ho and Wenwei Xu, "An Accurate Algebraic Solution for Moving Source Location Using TDOA and FDOA Measurements," *IEEE Transactions on Signal Processing*, vol. 52, pp. No.9, September 2004.

[39] Xiaoning Lu and K.C. Ho, "Analysis of The Degradation in Source Location Accuracy in The Presence of Sensor Location Error," *IEEE ICASSP*, 2006

[40] K.C. Ho, Xiaoning Lu and L. Kovavisaruch, "Source Localization Using TDOA and FDOA Measurements in the Presence of Receiver Location Errors: Analysis and Solution," *IEEE Transactions on Signal Processing*, vol. 55, pp. No.2, February 2007.