Localization Algorithms for Passive Sensor Networks

by

Darya Ismailova
B.Eng., University of Astrakhan, 2010

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF APPLIED SCIENCE

in the Department of Electricanl and Computer Engineering

Localization Algorithms for Passive Sensor Networks

by

Darya Ismailova
B.Eng., University of Astrakhan, 2010

Supervisory Committee

---

Dr. Wu-Sheng Lu, Supervisor
(Department of Electrical and Computer Engineering)

---

Dr. Pan Agathoklis, Departmental Member
(Department of Electrical and Computer Engineering)

**Supervisory Committee**

---

Dr. Wu-Sheng Lu, Supervisor
(Department of Electrical and Computer Engineering)

---

Dr. Pan Agathoklis, Departmental Member
(Department of Electrical and Computer Engineering)

## ABSTRACT

...

# Contents

# List of Abbreviations

| | |
|---|---|
| LS | Least Squares |
| ML | Maximum Likelihood |
| MDS | Multidimensional Scaling |
| DW-MDS | Distributed Weighted-MultidiDentional Scaling |
| SR-LS | |
| SRD-LS | |
| PDF | Probability Density Function |
| SPF | standard fixed point |
| SWLS | sequential weighted least squares |
| WSR-LS | weighted squared range based least squares (WSR-LS) |
| WSRD-LS | weighted squared range-difference based least squares (WSR-LS) |
| GTRS | |
| IRWSR-LS | |
| IRWSRD-LS | |
| MSE | |
| TDOA | |
| TOA | |
| WCDMA | |
| LTE | |
| O-TDOA | |
| CRLB | Cramér-Rao lower bound |
| NLLS | Non-Linear Least Squares |
| SMACOF | Scaling by MAjorizing a COmplicated Function |
| RSS | Received Signal Strength |
| NLOS | Non-Line Of Sight |
| UWB | Ultra Wide Band |
| SDP | SemiDefinite Programming |
| DC | Difference of Convex |
| PCCP | Penalty Convex Concave Procedure |
| CCP | Convex Concave Procedure |
| LP | Linear Program |

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 The Localization Problem

TODO

Review of ranging and localization methods, theory behind it, application, limitations.

TOA,

TDOA,

AOA ?,

non-range-based?

"Geolocation techniques"

[44] [28]

## 1.2 Contributions and Organization of the Thesis

### 1.2.1 Contributions of the Thesis

something like:

the material developed here / mathematical tools and methods are suitable for many different scenarios OR can have different world life applications, for example: TOA, TDOA, static positioning using UWB range measurements.

This work mostly investigated possible localization solution for the wireless sensor networks that use radio frequency measurements to obtain the range or range-difference measurements. However, some of the methods presented in this work can also be applied for other mediums, i.e. ultrasonic, sonic, light.

### 1.2.2 Organization of the Thesis

# Chapter 2

# Iterative Re-Weighting Least-Squares Methods for Source Localization

Locating a radiating source from range or range-difference measurements in a passive sensor network has recently attracted an increasing amount of research interest as it finds applications in a wide range of network-based wireless systems. Among the useful localization methods that have been documented over the years, least squares based algorithms constitute an important class of solution techniques as they are geometrically meaningful and often provide low complexity solution procedures with competitive estimation accuracy [1] - [15]. On the other hand, the error measure in a least squares (LS) formulation for the localization problem of interest is shown to be highly nonconvex, possessing multiple local solutions with degraded performance. There are many methods for continuous unconstrained optimization [42], however most of them are *local* methods that are sensitive to where the iteration begins, and give no guarantee to yield global solutions when applied to non-convex objective functions. In the case of source localization, this inherent feature of local methods is particular problematic because the source location is assumed to be entirely unknown and can appear practically anywhere, thus the chances to secure a good initial point for a local algorithm are next to none. For these reasons, various "global" localization techniques were investigated that are either non-iterative or insensitive to initial iterate. One representative in the class of global localization methods is the convex-relaxation based algorithm for range measurements proposed in [11], where

the least squares model is relaxed to a semidefinite programming problem which is known to be convex [40], hence robust to where it starts. Another representative in this class is reference [15], where localization problems for range as well as range difference measurements are addressed by developing solution methods for *squared* range LS (SR-LS) and *squared* range difference LS (SRD-LS) problems. The methods proposed in [15] are non-iterative and the solutions obtained are proven to be the global minimizers of the respective SR-LS and SRD-LS problems, which are shown to be excellent estimates of the original LS solutions.

This chapter presents improved least squares methods that demonstrate improved localization performance when compared with some best known results from the literature. The key new ingredient of the proposed algorithms is an iterative procedure where the SR-LS (SRD-LS) algorithm is iteratively applied to a weighted sum of squared terms where the weights are carefully designed so that the iterates produced quickly converge to a solution which is found to be considerably closer to the original range-based (range-difference-based) LS solution.

## 2.1 Source Localization Using Range Measurements

### 2.1.1 Problem Statement

The source localization problem considered here involves a given array of $m$ sensors specified by $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\}$ where $\boldsymbol{a}_i \in R^n$ contains the $n$ coordinates of the $i$th sensor in space $R^n$. Each sensor measures its distance to a radiating source $\boldsymbol{x} \in R^n$. Throughout it is assumed that only noisy copies of the distance data are available, hence the *range measurements* obey the model

$$r_i = \|\boldsymbol{x} - \boldsymbol{a}_i\| + \varepsilon_i, i = 1, \ldots, m. \tag{2.1}$$

where $\varepsilon_i$ denotes the unknown noise that has occurred when the $i$th sensor measures its distance to source $\boldsymbol{x}$. Let $\boldsymbol{r} = [r_1\ r_2 \ldots r_m]^T$ and $\boldsymbol{\varepsilon} = [\varepsilon_1\ \varepsilon_2 \ldots \varepsilon_m]^T$. The source localization problem can be stated as to estimate the exact source location $\boldsymbol{x}$ from the noisy range measurements $\boldsymbol{r}$. In the rest of this section, a least-squares (LS) formulation of the localization problem and two most relevant state-of-the-art solution methods are briefly reviewed; and a new method based on iterative re-weighting of

squared range LS technique as well as a variant of the proposed method are then presented.

### 2.1.2 LS Formulations and Review of Related Work

Least squares approaches have proven effective for source localization problems [1] -[15]. For the localization problem at hand, the range-based least squares (R-LS) estimate refers to the solution of the problem

$$\underset{\boldsymbol{x}}{\text{minimize}}\, f(\boldsymbol{x}) \equiv \sum_{i=1}^{m} (r_i - \|\boldsymbol{x} - \boldsymbol{a}_i\|)^2 \tag{2.2}$$

The primary reason that justifies formulation (2.2) is its connection to the maximum-likelihood location estimation that determines $\boldsymbol{x}$ by examining the probabilistic model of the error vector $\boldsymbol{\varepsilon}$. Assuming the errors $\varepsilon_i$ are independent and identically distributed (i.i.d.) Gaussian variables with zero mean and variance $\sigma_i^2$, then $\boldsymbol{\varepsilon}$ obeys a Gaussian distribution with zero mean and covariance $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2)$, and the maximum likelihood (ML) location estimator in this case is known to be

$$\boldsymbol{x}_{ML} = \arg\min_{\boldsymbol{x} \in R^n} (\boldsymbol{r} - \boldsymbol{g})^T \Sigma^{-1} (\boldsymbol{r} - \boldsymbol{g}) \tag{2.3}$$

where $\boldsymbol{g} = [g_1\ g_2 \ldots\ g_m]^T$ with

$$g_i = \|\boldsymbol{x} - \boldsymbol{a}_i\| \tag{2.4}$$

It follows immediately that the ML solution in (2.3) is identical to the R-LS solution of problem (2.2) when covariance $\boldsymbol{\Sigma}$ is proportional to the identity matrix, i.e., $\sigma_1^2 = \ldots = \sigma_m^2$. In the literature this is known as the equal noise power case. For notation simplicity the method described in this chapter focuses on the equal noise power case, however the method developed below is also applicable to the unequal noise power case by working on a weighted version of the objective in (2.2) with $\{\sigma_i^{-2}, i = 1, \ldots, m\}$ as the weights.

Although many methods for unconstrained optimization are available [42], most of them are *local* methods in the sense they are sensitive to the choice of initial point where the iteration of an optimization algorithm begins. Especially when applied to a nonconvex objective function which possesses a number of local minimizers, unless a chosen local method starts at an initial point that happens to be sufficiently close to the (unknown) global minimizer, the solution obtained by the method gives no

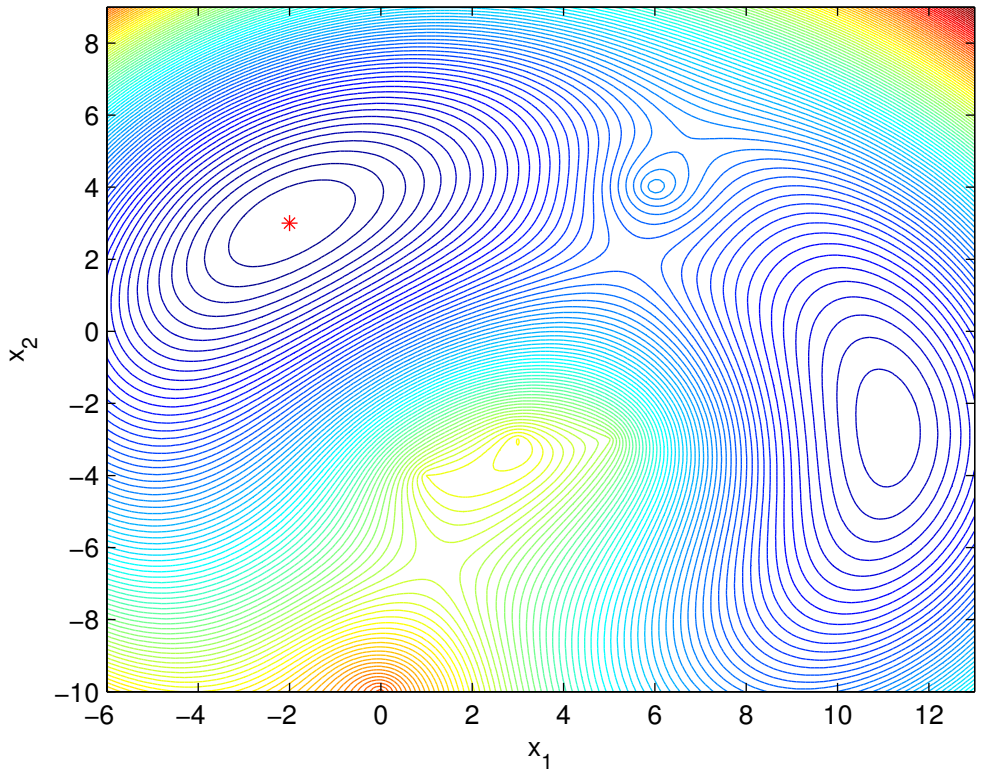



Figure 2.1: Contours of the R-LS objective function over the region $\Re = \{\boldsymbol{x} : -6 \leq x_1 \leq 13, -10 \leq x_2 \leq 9\}$

guaranty about global minimality. Unfortunately, the objective in (2.2) is highly nonconvex, possessing many local minimizers even for small-scale systems. As an example, consider an instance of the source localization problem on the plane $n = 2$ with five sensors $m = 5$ located at $(6, 4)^T, (0, -10)^T, (5, -3)^T, (1, -4)^T$ and $(3, -3)^T$ with the source emmitting the signal at $\boldsymbol{x}_s = (-2, 3)^T$. Figure 2.1 depicts a contour plot of the R-LS objective function in (2.2) over the region $\Re = \{\boldsymbol{x} : -6 \leq x_1 \leq 13, -10 \leq x_2 \leq 9\}$. It can be observed from the plot that there are two minimizers at $\tilde{\boldsymbol{x}} = (-1.9907, 3.0474)^T$ and $\widehat{\boldsymbol{x}} = (11.1152, -2.6785)^T$ with values of the objective $f(\tilde{\boldsymbol{x}}) = 0.1048$ and $f(\widehat{\boldsymbol{x}}) = 15.0083$ respectively. As expected, the global minimizer of R-LS objective offers a good approximation of the exact source location $\boldsymbol{x}_s$, but is unlikely to be precisely at point $\boldsymbol{x}_s$ because the objective $f(\boldsymbol{x})$ is defined using noisy range measurements. Note that for the exact source location $\boldsymbol{x}_s$ we have $f(\boldsymbol{x}_s) = \sum_{i=1}^{m} \varepsilon_i^2$.

Reference [11] addresses problem (2.2) by a convex relaxation technique where (2.2) is modified to a convex problem known as semidefinite programming (SDP) [40]. A key step in this procedure is to use (2.4) with $g_i$ as new variables, which leads (2.2) to the constrained problem

$$\underset{\boldsymbol{x},\boldsymbol{g}}{\text{minimize}} \sum_{i=1}^{m} (r_i - g_i)^2 \tag{2.5a}$$

$$\text{subject to:} \quad g_i^2 = \|\boldsymbol{x} - \boldsymbol{a}_i\|^2, \quad i = 1, \ldots, m. \tag{2.5b}$$

By further defining matrix variables

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{g} \\ 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{g}^T & 1 \end{bmatrix} \text{ and } \boldsymbol{X} = \begin{bmatrix} \boldsymbol{x} \\ 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^T & 1 \end{bmatrix} \tag{2.6}$$

and neglecting the rank constrains on $\boldsymbol{G}$ and $\boldsymbol{X}$, (2.5) can be reformulated in term of variables $\boldsymbol{G}$ and $\boldsymbol{X}$ as

$$\underset{\boldsymbol{X},\boldsymbol{G}}{\text{minimize}} \sum_{i=1}^{m} \left( G_{ii} - 2r_i G_{m+1,i} + r_i^2 \right) \tag{2.7a}$$

$$\text{subject to:} \quad G_{ii} = Tr\left(\boldsymbol{C}_i \boldsymbol{X}\right), i = 1, \ldots, m \tag{2.7b}$$

$$\boldsymbol{G} \succeq 0, \boldsymbol{X} \succeq 0 \tag{2.7c}$$

$$G_{m+1,m+1} = G_{n+1,n+1} = 1 \tag{2.7d}$$

where

$$C_i = \begin{pmatrix} \boldsymbol{I}_{n\times n} & -\boldsymbol{a}_i \\ -\boldsymbol{a}_i^T & \|\boldsymbol{a}_i\|^2 \end{pmatrix} \quad i = 1, \ldots, m \tag{2.8}$$

which is a standard SDP problem that can be solved efficiently [40, 42]. Note that because (2.7) is a convex problem, global minimality of the solution is ensured regardless of the initial point used. On the other hand, however, because (2.7) is an approximation of the original problem in (2.2), the solution of (2.7) is only an approximate solution of problem (2.2). In what follows the solutions obtained by this SDP-relaxation based method will be referred to as SDR-LS solutions.

A rather different approach is recently proposed in [15] where the localization problem (2.2) is tackled by developing techniques that find global solution of the

*squared range based LS* (SR-LS) problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \sum_{i=1}^{m} \left( \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 - r_i^2 \right)^2 \tag{2.9}$$

By writing the objective in (2.9) as $\left(\alpha - 2\boldsymbol{a}_i^T\boldsymbol{x} + \|\boldsymbol{a}_i\|^2 - r_i^2\right)^2$ with $\alpha = \|\boldsymbol{x}\|^2$, it becomes a convex quadratic objective if one treats $\alpha$ as an additional variable and $\alpha = \|\boldsymbol{x}\|^2$ as a constraint. In this way, (2.9) is converted to the following constrained LS problem after necessary variable changes:

$$\underset{\boldsymbol{y} \in R^{n+1}}{\text{minimize}} \|\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}\|^2 \tag{2.10a}$$

$$\text{subject to:} \quad \boldsymbol{y}^T\boldsymbol{D}\boldsymbol{y} + 2\boldsymbol{f}^T\boldsymbol{y} = 0 \tag{2.10b}$$

where

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{x} \\ \|\boldsymbol{x}\|^2 \end{pmatrix}, \ \boldsymbol{A} = \begin{pmatrix} -2\boldsymbol{a}_1^T & 1 \\ \vdots & \vdots \\ -2\boldsymbol{a}_m^T & 1 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} r_1^2 - \|\boldsymbol{a}_1\|^2 \\ \vdots \\ r_m^2 - \|\boldsymbol{a}_m\|^2 \end{pmatrix} \tag{2.11}$$

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{I}_{n\times n} & \boldsymbol{0}_{n\times n} \\ \boldsymbol{0}_{n\times n} & 0 \end{pmatrix}, \ \boldsymbol{f} = \begin{pmatrix} \boldsymbol{0} \\ -0.5 \end{pmatrix}$$

This problem conversion, made in [15], turns out to be crucial as problem (2.10), which remains to be nonconvex because of the nonlinear equality constraint (2.10b), falls into the class of generalized trust region subproblems (GTRS) [20, 21] whose global solutions can be computed by exploring the KKT conditions which are both necessary and sufficient optimality conditions in this case [20].

We now conclude this section with a couple of remarks. First, an unconstrained version of (2.10) may be obtained by neglecting the constraint in (2.10b) as

$$\underset{\boldsymbol{y} \in R^{n+1}}{\text{minimize}} \|\boldsymbol{A}\boldsymbol{y} - \boldsymbol{b}\|^2 \tag{2.12}$$

whose solution, called *unconstrained squared-range-based LS* (USR-LS) estimate, is given by

$$\boldsymbol{y}^* = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{b} \tag{2.13}$$

It is demonstrated by numerical experiments [15] that the SR-LS solution outperforms

the USR-LS and, in many cases, SDR solutions. Second, the SR-LS solution, although it solves (2.9) exactly, lacks the statistical interpretation of the ML formulation. The SR-LS remains to be an approximate solution for the original problem in (2.2) and, as it was demonstrated by the numerical results in [16] and [18], provides less accurate estimates of the true source location, than the LS estimate. The method, described in detail below, tries to reduce the gap between the two solutions.

### 2.1.3 An Iterative Re-Weighting Approach

Iterative re-weighting least squares method is a popular technique used for solving problems involving the sums of norms. The method has found many applications, such as in robust regression [22, 23], sparse recovery [17], but the most relevand application for the current case is for solving the Fermat-Weber location problem. The Fermat–Weber problem has a long history and has been extensively studied in the field of optimization and location theory [23]. This problem can be stated as: Given $m$ points $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m \in R^n$ called *anchors* and nonnegative weights $\omega_1, \omega_2, \ldots, \omega_m > 0$, find $\boldsymbol{x} \in R^n$ that minimizes the weighted sum of Euclidian distances between $\boldsymbol{x}$ and the $m$ anchors:

$$\underset{\boldsymbol{x} \in R^n}{\text{minimize}} \sum_{i=1}^{m} \omega_i \|\boldsymbol{x} - \boldsymbol{a}_i\|.$$

Fermat–Weber problem is much easier to analyze and solve than the ML problem (2.2) because it is a well-structured nonsmooth convex minimization problem. The similarities between the Fermat–Weber problem and problem (2.2) have been noted and addressed in the literature [16] with a gradient method with a fixed step size, known as the standard fixed point (SFP) algorithm, to deal with problem (2.2). However, being a gradient method, likelihood for the SFP algorithm to converge to a local solution exists. Another method, also proposed in [16] and known as the the sequential weighted least squares algorithm (SWLS), is also an iterative method where each iteration involves solving a nonlinear least squares problem similar to (2.9). The SWLS method is found to be superior over SFP in terms of convergence rate and a wider region of convergence to the global minimum [16]. However, the possibility for SWLS to converge to a local minimum remains in certain sensor setups even if the initial point is constructed using a procedure developed specifically for SWLS. The method presented below takes an approach that is different from those described above in the sense that it does not require an initial point and the solution produced

is guranteed to converge to a *global* solution.

**Weighted squared range based least squares formulation**

We now consider the weighted squared range based least squares (WSR-LS) problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \sum_{i=1}^{m} w_i \left( \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 - r_i^2 \right)^2 \tag{2.14}$$

which is obviously a weighted version of the SR-LS problem in (2.9). Following [15], it is rather straightforward to convert (2.14) into a GTRS similar to (2.10) as

$$\underset{\boldsymbol{y} \in R^{n+1}}{\text{minimize}} \, \|\boldsymbol{\Gamma} \left( \boldsymbol{A}\boldsymbol{y} - \boldsymbol{b} \right) \|^2 \tag{2.15a}$$

$$\text{subject to:} \quad \boldsymbol{y}^T \boldsymbol{D} \boldsymbol{y} + 2\boldsymbol{f}^T \boldsymbol{y} = 0 \tag{2.15b}$$

where $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{D}$, and $\boldsymbol{f}$ are defined in (2.11) and $\boldsymbol{\Gamma} = \text{diag}\left(\sqrt{w_1}, \ldots, \sqrt{w_m}\right)$. Clearly, problem (2.15) can be written as

$$\underset{\boldsymbol{y} \in R^{n+1}}{\text{minimize}} \, \|\boldsymbol{A}_w \boldsymbol{y} - \boldsymbol{b}_w\|^2 \tag{2.16a}$$

$$\text{subject to:} \quad \boldsymbol{y}^T \boldsymbol{D} \boldsymbol{y} + 2\boldsymbol{f}^T \boldsymbol{y} = 0 \tag{2.16b}$$

where $\boldsymbol{A}_w = \boldsymbol{\Gamma}\boldsymbol{A}$ and $\boldsymbol{b}_w = \boldsymbol{\Gamma}\boldsymbol{b}$. On comparing (2.16) with (2.10), if $S(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{D}, \boldsymbol{f})$ denotes a solver that produces the global solution of problem (2.10) for a given data set $\{\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{D}, \boldsymbol{f}\}$, then the same solver produces the global solution of the weighted problem (2.14) as long as it is applied to the data set $\{\boldsymbol{A}_w, \boldsymbol{b}_w, \boldsymbol{D}, \boldsymbol{f}\}$. We stress that the weights $\{w_i, i = 1, \ldots, m\}$ in (2.14) are *fixed* nonnegative constants.

**Moving the SR-LS solution towards R-LS solution via iterative re-weighting**

The main idea here is to use the weights $\{w_i, i = 1, \ldots, m\}$ to tune the objective in (2.14) toward the objective in (2.2) so that the solution obtained by minimizing such a WSR-LS objective is expected to be closer toward that of the problem (2.2). To substantiate the idea, we compare the $i$th term of the objective in (2.14) with its counterpart in (2.2), namely,

$$\underbrace{w_i \left( \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 - r_i^2 \right)^2}_{\text{in (15)}} \leftrightarrow \underbrace{\left( \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i \right)^2}_{\text{in (2)}} \tag{2.17}$$

and write the term in (2.14) as

$$\begin{aligned} & w_i \left( \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 - r_i^2 \right)^2 = \\ & w_i \left( \|\boldsymbol{x} - \boldsymbol{a}_i\| + r_i \right)^2 \underbrace{\left( \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i \right)^2}_{\text{same as in (2)}} \end{aligned}$$

It follows that the objective in (2.14) would be the same as in (2.2) if the weights $w_i$ were assigned to $1/\left( \|\boldsymbol{x} - \boldsymbol{a}_i\| + r_i \right)^2$. Evidently, weight assignments as such cannot be realized because $w_i$'s must be fixed constants for (2.14) to be a globally solvable WSR-LS problem. A natural remedy to deal with this technical difficulty is to employ an iterative procedure whose $k$th iteration generates a global solution $\boldsymbol{x}_k$ of a WSR-LS sub-problem of the form

$$\underset{x}{\text{minimize}} \sum_{i=1}^{m} w_i^{(k)} \left( \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 - r_i^2 \right)^2 \tag{2.18}$$

where for $k \geq 2$ the weights $\{w_i^{(k)}, i = 1, \ldots, m\}$ are assigned using the previous iterate $\boldsymbol{x}_{k-1}$ as

$$w_i^{(k)} = \frac{1}{\left( \|\boldsymbol{x}_{k-1} - \boldsymbol{a}_i\| + r_i \right)^2} \tag{2.19}$$

while for $k = 1$ all weights $\{w_i^{(1)}, i = 1, \ldots, m\}$ are set to unity. Clearly the weights given by (2.19) are realizable. More importantly, when the iterates produced by solving (2.18) (namely $\boldsymbol{x}_k$ for $k = 1, 2, \ldots$) converge, in the $k$th iteration with $k$ sufficiently large, the objective function of (2.18) in a small vicinity of its solution $\boldsymbol{x}_k$ is approximately equal to

$$\begin{aligned} & \sum_{i=1}^{m} w_i^{(k)} \left( \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 - r_i^2 \right)^2 \\ \approx & \sum_{i=1}^{m} w_i^{(k)} \left( \|\boldsymbol{x}_k - \boldsymbol{a}_i\|^2 - r_i^2 \right)^2 \\ = & \sum_{i=1}^{m} w_i^{(k)} \left( \|\boldsymbol{x}_k - \boldsymbol{a}_i\| + r_i \right)^2 \left( \|\boldsymbol{x}_k - \boldsymbol{a}_i\| - r_i \right)^2 \\ \approx & \sum_{i=1}^{m} w_i^{(k)} \left( \|\boldsymbol{x}_{k-1} - \boldsymbol{a}_i\| + r_i \right)^2 \left( \|\boldsymbol{x}_k - \boldsymbol{a}_i\| - r_i \right)^2 \\ = & \sum_{i=1}^{m} \left( \|\boldsymbol{x}_k - \boldsymbol{a}_i\| - r_i \right)^2 \approx \sum_{i=1}^{m} \left( \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i \right)^2 \end{aligned}$$

In words, with the weights from (2.19), the limiting point of the iterates produced by iteratively solving WSR-LS problem (2.18) is expected to be close to the global solution of problem (2.2).

The algorithmic steps of the proposed localization method are outlined as follows.

**Algorithm 1**

1) Input data: Sensor locations $\{\boldsymbol{a}_i, i = 1, \ldots, m\}$, range measurements $\{r_i, i = 1, \ldots, m\}$, maximum number of iterations $k_{max}$ and convergence tolerance $\zeta$.

2) Generate data set $\{\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{d}, \boldsymbol{f}\}$ as

$$\boldsymbol{A} = \begin{pmatrix} -2\boldsymbol{a}_1^T & 1 \\ \vdots & \vdots \\ -2\boldsymbol{a}_m^T & 1 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} r_1^T - \|\boldsymbol{a}_1\|^T \\ \vdots \\ r_m^T - \|\boldsymbol{a}_m\|^T \end{pmatrix}$$

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{I}_{n\times n} & \boldsymbol{0}_{n\times n} \\ \boldsymbol{0}_{n\times n} & 0 \end{pmatrix}, \boldsymbol{f} = \begin{pmatrix} \boldsymbol{0} \\ -0.5 \end{pmatrix}.$$

Set $k = 1, w_i^{(1)} = 1$ for $i = 1, \ldots, m$.

3) Set $\boldsymbol{\Gamma}_k = \text{diag}\left(\sqrt{w_1^{(k)}}, \ldots, \sqrt{w_m^{(k)}}\right)$, $\boldsymbol{A}_w = \boldsymbol{\Gamma}_k \boldsymbol{A}$ and $\boldsymbol{b}_w = \boldsymbol{\Gamma}_k \boldsymbol{b}$.

4) Solve the WSR-LS problem

$$\underset{x}{\text{minimize}} \sum_{i=1}^{m} w_i^{(k)} \left(\|\boldsymbol{x} - \boldsymbol{a}_i\|^2 - r_i^2\right)^2$$

by solving (2.16), i.e.

$$\underset{\boldsymbol{y}\in R^{n+1}}{\text{minimize}} \|\boldsymbol{A}_w \boldsymbol{y} - \boldsymbol{b}_w\|^2$$
$$\text{subject to: } \boldsymbol{y}^T \boldsymbol{D} \boldsymbol{y} + 2\boldsymbol{f}^T \boldsymbol{y} = 0$$

to obtain its global solution $\boldsymbol{x}_k$.

5) If $k = k_{max}$ or $\|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}\| < \zeta$, terminate and output $\boldsymbol{x}_k$ as the solution; otherwise, set $k = k + 1$, update weights $\{w_i^{(k)}, i = 1, \ldots, m\}$ using

$$w_i^{(k)} = \frac{1}{\left(\|\boldsymbol{x}_{k-1} - \boldsymbol{a}_i\| + r_i\right)^2}$$

and repeat from Step 3).

From the steps in Algorithm 1, it follows that the complexity of the algorithm is practically equal to the complexity of the WSR-LS solver involved in Step 4 times the number of iterations, $k$. Computer simulations have indicated that the algorithm converges with a small number of iterations, typically a $k \leq 6$ suffices. For simplicity, the solutions obtained from Algorithm 1 are called IRWSR-LS solutions. Technical details on how to solve (2.16) can be found in Appendix 1.

**A variant of Algorithm 1**

As argued above, the IRWSR-LS solution from Algorithm 1 is expected to be an improved approximation of the global solution of R-LS problem in (2.2). However a small gap between the two solutions is inevitable owing to the approximate nature of the re-weighting strategy. In what follows we present a variant of Algorithm 1 that closes this gap by taking the IRWSR-LS solution as an initial point to run a good local unconstrained optimization algorithm for problem (2.2). The rationale behind this two-step approach is that the initial point produced in the first step by Algorithm 1 is highly likely within a sufficiently small vicinity of the exact global solution of problem (2.2), hence a good local method will lead it to the exact solution in a small number of iterations. We remark that such a "hybrid" approach is also expected to work with other "global" methods including the SDR-LS and SR-LS methods, but with a difference that employing an IRWSR-LS solution in the first step improves the closeness of the initial point, hence increases the likelihood of securing the exact global solution of problem (2.2) by a local method in the second step.

For the localization problem in question, the well-known Newton algorithm [42] is chosen as a local method because of its fast convergence and low complexity. We note that, unlike in a general scenario where the Newton algorithm is often considered numerically expensive because it requires to compute the inverse of the Hessian matrix, computing such an inverse is not costly in the present case because the dimension of the unknown vector $\boldsymbol{x}$ is extremely low: $n = 2$ or 3. Moreover, the Hessian matrix involved can be efficiently evaluated by a closed-form formula, as shown below.

To evaluate the Hessian of the objective $f(\boldsymbol{x})$ in (2.2), we assume $\boldsymbol{x} \neq \boldsymbol{a}_i$ for $i = 1, \ldots, m$, so that $f(\boldsymbol{x})$ is a smooth function of $\boldsymbol{x}$. The assumption simply means that the radiating source is away from the sensor at least by a certain distance, which appears to be reasonable for a practical localization problem. Under this

circumstance, the gradient and Hessian of $f(\boldsymbol{x})$ are found to be

$$\boldsymbol{g}(\boldsymbol{x}) = \sum_{i=1}^{m} \left(1 - \frac{r_i}{\|\boldsymbol{x} - \boldsymbol{a}_i\|})\right) (\boldsymbol{x} - \boldsymbol{a}_i) \tag{2.20a}$$

and

$$\boldsymbol{H}(\boldsymbol{x}) = 2\left(\tau \boldsymbol{I} + \boldsymbol{H}_1(\boldsymbol{x})\right) \tag{2.20b}$$

respectively, where

$$\tau = m - \sum_{i=1}^{m} \frac{r_i}{\|\boldsymbol{x} - \boldsymbol{a}_i\|}$$

and

$$\boldsymbol{H}_1(\boldsymbol{x}) = \sum_{i=1}^{m} \frac{r_i (\boldsymbol{x} - \boldsymbol{a}_i)(\boldsymbol{x} - \boldsymbol{a}_i)^T}{\|\boldsymbol{x} - \boldsymbol{a}_i\|^3}.$$

To ensure a descent Newton step, the positive definiteness of the Hessian $\boldsymbol{H}(\boldsymbol{x})$ needs to be examined and, in case $\boldsymbol{H}(\boldsymbol{x})$ is not positive definite, to be modified to guatrantee its positive definiteness. To this end, the eigen-decomposition of $\boldsymbol{H}(\boldsymbol{x})$, namely,

$$\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$$

may be performed, where $\boldsymbol{U}$ is orthogonal and $\boldsymbol{\Lambda} = \text{diag}\,(\tau + \ \lambda_1, \ldots, \tau + \lambda_n)$ with $\{\lambda_i, i = 1, \ldots, n\}$ being eigenvalues of $\boldsymbol{H}_1(\boldsymbol{x})$. Let $l_{min}$ be the smallest eigenvalue of $\boldsymbol{H}(\boldsymbol{x})$, namely $l_{min} = \min(\tau + \lambda_1, \ldots, \tau + \lambda_n)$. If $l_{min} > 0$, then $\boldsymbol{H}(\boldsymbol{x})$ is positive definite and the Newton algorithm is carried out without modification; if $l_{min} \leq 0$, then the algorithm uses a slightly modified Hessian given by

$$\tilde{\boldsymbol{H}}(\boldsymbol{x}) = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}\boldsymbol{U}^T$$

where $\tilde{\boldsymbol{\Lambda}} = \text{diag}\left(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n\right)$

$$\tilde{\lambda}_i = \begin{cases} \tau + \lambda_i & \text{if } \tau + \lambda_i > 0 \\ \delta & \text{if } \tau + \lambda_i \leq 0 \quad i = 1, \ldots, m \end{cases}$$

and $\delta$ a small positive constant. Obviously, $\tilde{\boldsymbol{H}}(\boldsymbol{x})$ is guaranteed to be positive definite. The search direction in the $k$th iteration of the modified Newton algorithm is given by

$$\boldsymbol{d}_k = -\boldsymbol{U}\tilde{\boldsymbol{\Lambda}}^{-1}\boldsymbol{U}^T g(\boldsymbol{x}_k)$$

where $g(\boldsymbol{x}_k)$ is given by (2.20). In what follows, solutions obtained by the proposed two-step method are called *hybrid* IRWSR-LS solutions.

### 2.1.4 Numerical Results

Performance of the proposed algorithms was evaluated and compared with existing state-of-the-art methods by Monte Carlo simulations with a set-up similar to that of [15]. SR-LS solutions were used as performance benchmark for Algorithm 1 and its variant. Our simulation studies of Algorithm 1 and its variant considered a scenario that consists of $m = 5$ sensors $\{\boldsymbol{a}_i,\ i = 1, 2, \ldots, m\}$ randomly placed in the planar region in $[-15; 15] \times [-15; 15]$, and a radiating source $\boldsymbol{x}_s$, located randomly in the region $[-10; 10] \times [-10; 10]$. Coordinates of the source and sensors were generated for each dimension following a uniform distribution. The range measurements $\{r_i,\ i = 1, 2, \ldots, m\}$ were calculated using (2.1) and Step 4 of Algorithm 1 was implemented using the SR-LS algorithm proposed in [15]. Measurement noise $\{\varepsilon_i,\ i = 1, \ldots, m\}$ was modelled as i.i.d. Gaussian random variables with zero mean and variance $\sigma^2$, with $\sigma$ being one of three possible levels $\{10^{-3}, 10^{-2}, 10^{-1}\}$. Accuracy of source location estimation was evaluated as the average of the squared position error $\|\boldsymbol{x}^* - \boldsymbol{x}_s\|^2$ where $\boldsymbol{x}_s$ denotes the exact source location and $\boldsymbol{x}^*$ is its estimation obtained by SR-LS, IRWSR-LS and hybrid-IRWSR-LS methods, respectively. Table 2.1 provides comparisons of these methods with SR-LS, where each table entry is a MSE averaged over 1,000 Monte Carlo runs of a given method for a given noise level. For the columns representing performance of the IRWSR-LS and *hybrid* IRWSR-LS methods each table entry lists their MSE and relative improvement over SR-LS solutions in percentage, in the format of MSE (% Improvement).

Table 2.1: MSE of position estimation for SR-LS, IRWSR-LS and *hybrid* IRWSR-LS methods

| $\sigma$ | SR - LS | IRWSR-LS (Im.,%) | *hybrid* IRWSR-LS (Im.,%) |
|---|---|---|---|
| 1e-03 | 2.03251062e-06 | 1.19962894e-06 (41) | 1.19949340e-06 (41) |
| 1e-02 | 1.83717590e-04 | 1.24797437e-04 (32) | 1.24812091e-04 (32) |
| 1e-01 | 1.83611315e-02 | 1.22233840e-02 (33) | 1.22139427e-02 (33) |

It is observed that IRWSR-LS solutions offer considerable improvement over SR-LS solutions, and, as expected, in most cases hybrid IRWSR-LS solutions provide

further but only incremental improvement. This is not surprising because the IRWSR-LS solutions themselves are already fairly close to the solutions of problem (2.2). It should also be noted again that for the exact source location $\boldsymbol{x}_s$ we have $f(\boldsymbol{x}_s) = \sum_{i=1}^{m} \varepsilon_i^2$. One might argue that the SR-LS solution already provides a rather good approximation for R-LS in the sense that SR-LS and IRWSR-LS (hybrid IRWSR-LS) have the same order of magnitude of the mean squared error. However, further analysis of the data that was used to generate Table 2.1 illustrates the advantage of the IRWSR-LS (hybrid IRWSR-LS) solution over the SR-LS.

Each entry in Table 2.2 is a standard deviation of the squared estimation errors aggregated over the same 1,000 Monte Carlo runs described above in Table 2.1 (where the MSE of the position estimation are shown). The results summarised in Table 2.2 demonstrate again that IRWSR-LS and hybrid IRWSR-LS outperform SR-LS. Figures 2.2 to 2.4 depict the histograms of the location estimation errors $\|\boldsymbol{x}^* - \boldsymbol{x}_s\|$ of the SR-LS solution (left images) and IRWSR-LS (right images) for all three noise levels with $\sigma$ being one of $\{10^{-3}, 10^{-2}, 10^{-1}\}$, where $\boldsymbol{x}^*$ denotes the estimated location and $\boldsymbol{x}_s$ is the exact location of the source. Note that the histograms that correspond to the results obtained by IRWSR-LS are shifted closer towards 0 than those obtained by SR-LS and have smaller variance, and the solutions obtained by running IRWSR-LS have fewer outliers.

Table 2.2: Standard deviation of the squared estimation error for SR-LS, IRWSR-LS and *hybrid* IRWSR-LS methods

| $\sigma$ | SR - LS | IRWSR-LS | *hybrid* IRWSR-LS |
|---|---|---|---|
| 1e-03 | 6.3438e-06 | 2.0843e-06 | 2.0864e-06 |
| 1e-02 | 3.2575e-04 | 2.0530e-04 | 2.0530e-04 |
| 1e-01 | 4.6998e-02 | 2.1377e-02 | 2.1377e-02 |

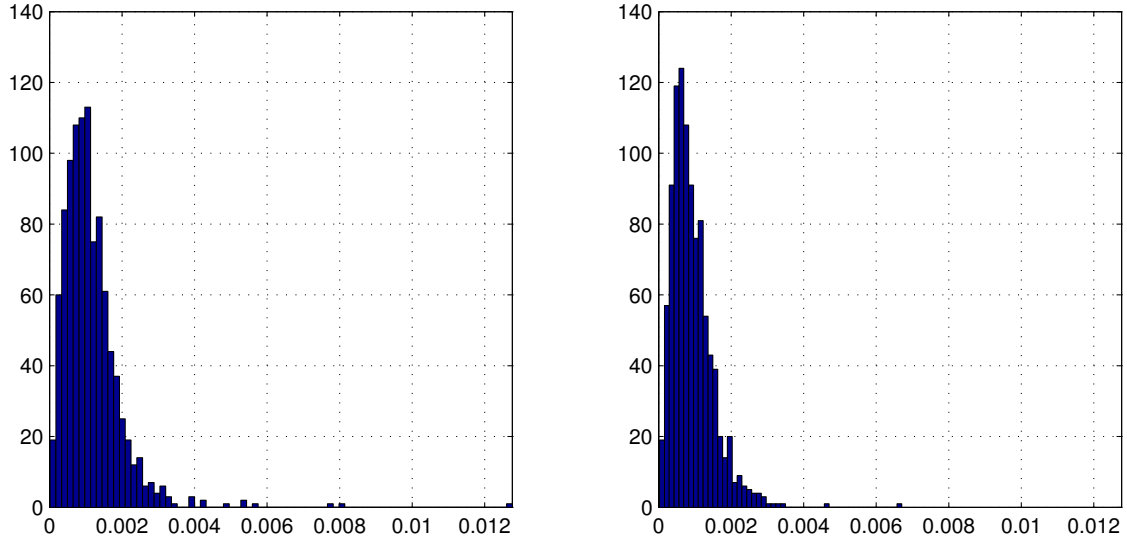



Figure 2.2: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 10^{-3}$

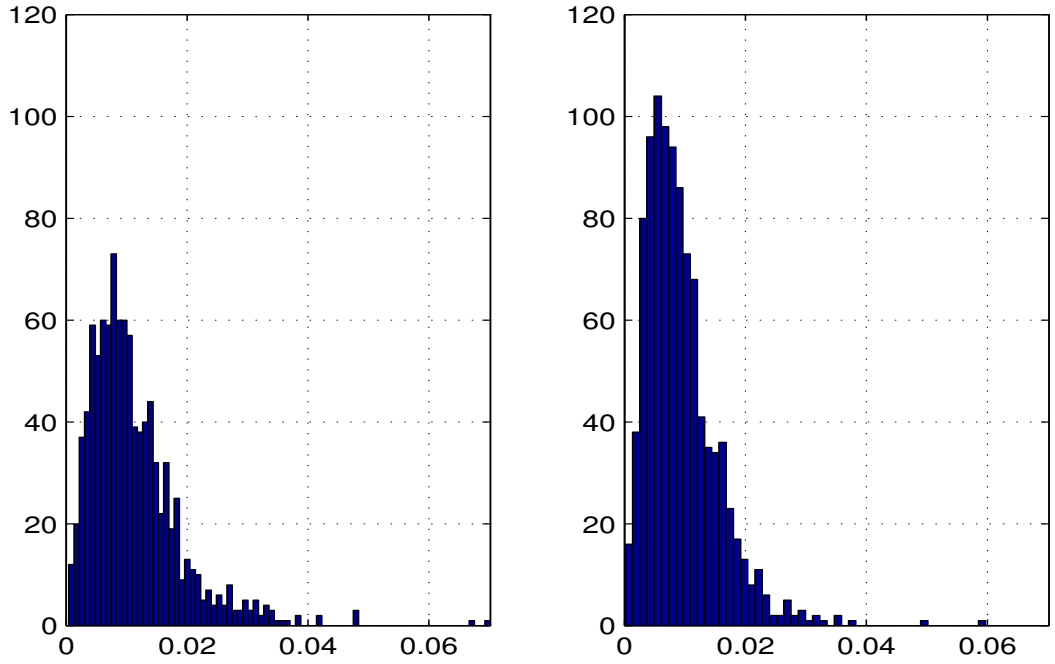



Figure 2.3: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 10^{-2}$

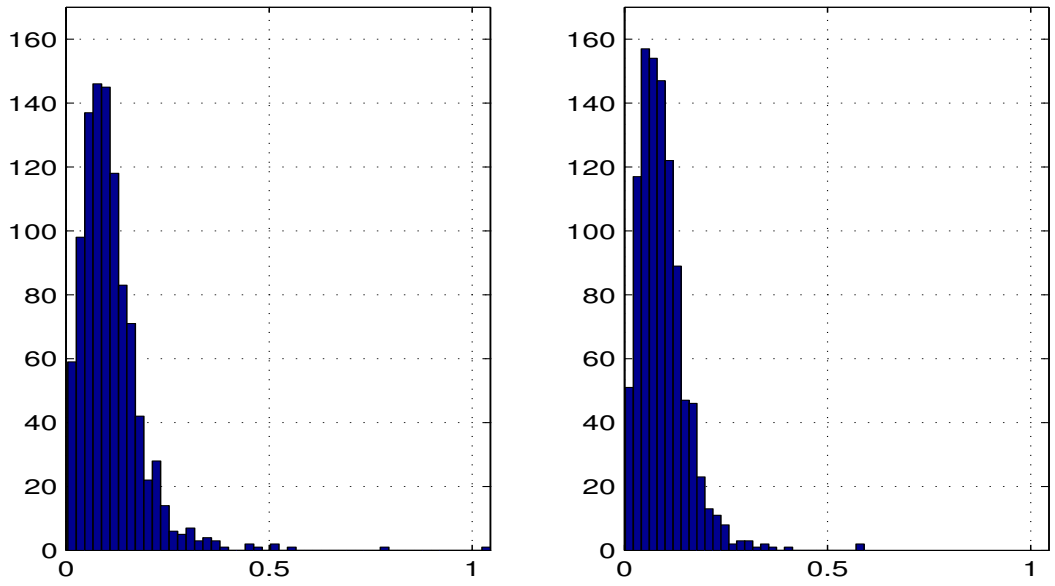



Figure 2.4: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 10^{-1}$

## 2.2 Source Localization Using Range-Difference Measurements

### 2.2.1 Problem Formulation

Another type of source localization problem that has attracted considerable attention is that of localizing a radiating source using range-difference measurements [14, 15]. In practice, range-difference measurements may be obtained from the time differences of arrival measured by an array of passive sensors. Time difference of arrival (TDOA) is a time-based positioning method based on the idea that the location of an active mobile unit (source of the signal) can be determined by examining the difference in time at which the signal arrives at multiple reference points. Adopting this technique is useful in practical scenarios where synchronization between mobile units is not available [44]. A typical example can be found in 3G (WCDMA) and LTE networks where the observed time difference of arrival (O-TDOA) technique is used to estimate the location of mobile units. We remark that in WCDMA networks, only the base stations are synchronized with each other, but mobile units are unsynchronized with base stations.

Each TDOA measurement constrains the location of the signal source to be on a hyperboloid with a constant range-difference between the two reference points. Specifically a TDOA measurement between base stations $BS_i$ and $BS_0$ is given by

$$t_{i0} = (t_i - t_x) - (t_0 - t_x) = t_i - t_0$$

where $t_x$ is the clock time of the mobile unit, $t_i$ and $t_0$ are the time of arrival between the mobile unit and stations $BS_i$ and $BS_0$ respectively. The above equation can be written in terms of distance (range-difference) through scaling

$$r_i = (t_i - t_0)c = r_i - r_0 = \|\boldsymbol{a}_i - \boldsymbol{x}\| - \|\boldsymbol{a}_0 - \boldsymbol{x}\|, i = 1, \ldots, m$$

where $c$ is the speed of signal propagation, $r_i$ is the distance from station $\boldsymbol{a}_i$ to source $\boldsymbol{x}$, $r_0$ is the distance from station $\boldsymbol{a}_0$ to source $\boldsymbol{x}$, and $\boldsymbol{a}_i \in R^n$ with $n = 2$ or 3, contains coordinates of the $i$th base station. Without loss of generality, the latter equation is valid with the assumption that the station $BS_0$ is placed at the origin of the coordinate system, i.e. $\boldsymbol{a}_0 = \boldsymbol{0}$ and used as a *reference* station [44].

The localization problem here is to estimate the location of a radiating source $\boldsymbol{x}$

given the locations of the $m+1$ sensors $\{\boldsymbol{a}_i, i = 0, 1, \ldots, m\}$ and noise-contaminated range-difference measurements $\{d_i, i = 1, 2, \ldots, m\}$ where

$$d_i = r_i + \varepsilon_i = \|\boldsymbol{a}_i - \boldsymbol{x}\| - \|\boldsymbol{x}\| + \varepsilon_i, \text{ for } i = 1, 2, \ldots, m \tag{2.21}$$

Therefore, the standard range-difference LS (RD-LS) problem is formulated as

$$\underset{\boldsymbol{x} \in R^n}{\text{minimize}}\, F(\boldsymbol{x}) = \sum_{i=1}^{m} \left(d_i + \|\boldsymbol{x}\| - \|\boldsymbol{x} - \boldsymbol{a}_i\|\right)^2 \tag{2.22}$$

Unfortunately, problem (2.22) is nonconvex and finding the global solution of (2.22) turns out to be a very hard problem. Nonlinear least squares (NLLS) algorithm [30] is widely used in TDOA localization systems for its performance. If the range measurement errors can be modeled as an additive white Gaussian noise, the accuracy of NLLS solution approaches the Cramér-Rao lower bound (CRLB). However, NLLS is not guaranteed to converge [13], [30] if the initial point is chosen far away from the actual source location. This becomes a more serious problem when the system coverage area is large since it becomes more difficult to secure an initial point that is close enough to the *unknown* source location. The Scaling by MAjorizing a COmplicated Function (SMACOF) strategy proposed in [31] can also be applied for position estimation. Compared with NLLS, it is not sensitive to the choice of the initial point and the mean-square error is guaranteed to decrease at each iteration, however the algorithm converges significantly slower. Reference [15] proposes a squared range-difference least squares (SRD-LS) approach to address this problem, which is summarized below.

By writing (2.21) as $d_i + \|\boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{a}_i\|$ and squaring both sides, we obtain

$$\left(d_i + \|\boldsymbol{x}\|\right)^2 = \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 \tag{2.23}$$

which can be simplified to

$$-2d_i\|\boldsymbol{x}\| - 2\boldsymbol{a}_i^T\boldsymbol{x} = g_i, i = 1, \ldots, m \tag{2.24}$$

where $g_i = d_i^2 - \|\boldsymbol{a}_i\|^2$. In practice (2.24) does not hold exactly due to measurement noise that contaminates the data $d_i$'s. In other words, if $d_i$'s in (2.24) are taken to be

real-world data, then we only have

$$-2d_i\|\boldsymbol{x}\| - 2\boldsymbol{a}_i^T\boldsymbol{x} - g_i \approx 0, i = 1, \ldots, m \tag{2.25}$$

Reference [15] proposes a LS solution for the problem at hand by minimizing the sum of squared residues on the left side of (2.25), namely,

$$\underset{\boldsymbol{x}\in R^n}{\text{minimize}} \sum_{i=1}^{m} \left(-2\boldsymbol{a}_i^T\boldsymbol{x} - 2d_i\|\boldsymbol{x}\| - g_i\right)^2 \tag{2.26}$$

By introducing new variable $\boldsymbol{y} = [\boldsymbol{x}^T \|\boldsymbol{x}\|]^T$ and noticing nonnegativity of the component $y_{n+1}$ problem (2.26) is converted to

$$\underset{\boldsymbol{y}\in R^{n+1}}{\text{minimize}} \|\boldsymbol{B}\boldsymbol{y} - \boldsymbol{g}\|^2 \tag{2.27a}$$

$$\text{subject to: } \boldsymbol{y}^T\boldsymbol{C}\boldsymbol{y} = 0 \tag{2.27b}$$

$$y_{n+1} \geq 0 \tag{2.27c}$$

where

$$\boldsymbol{g} = \begin{pmatrix} g_1 \\ \vdots \\ g_m \end{pmatrix}, \ \boldsymbol{B} = \begin{pmatrix} -2\boldsymbol{a}_1^T & -2d_1 \\ \vdots & \vdots \\ -2\boldsymbol{a}_m^T & -2d_m \end{pmatrix}, \ \boldsymbol{C} = \begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{0}_{n\times 1} \\ \boldsymbol{0}_{1\times n} & -1 \end{pmatrix} \tag{2.28}$$

Because of the presence of the nonnegativity constraint in (2.27c), (2.27) is no longer a GTRS problem hence the technique used for the case of range measurements does not apply. Nevertheless reference [15] presents a rigorous argument which shows that the optimal solution of (2.27) either assumes the form of

$$\tilde{\boldsymbol{y}}(\lambda) = \left(\boldsymbol{B}^T\boldsymbol{B} + \lambda\boldsymbol{C}\right)^{-1}\boldsymbol{B}^T\boldsymbol{g}$$

where $\lambda$ solves

$$\tilde{\boldsymbol{y}}(\lambda)^T\boldsymbol{C}\tilde{\boldsymbol{y}}(\lambda) = 0 \tag{2.29}$$

and makes $\boldsymbol{B}^T\boldsymbol{B} + \lambda\boldsymbol{C}$ positive definite, or is the vector among $\{\boldsymbol{0}, \tilde{\boldsymbol{y}}(\lambda_1), \ldots, \tilde{\boldsymbol{y}}(\lambda_p)\}$ that gives the smallest objective function in (2.27a), where $\{\lambda_i, i = 1, \ldots, p\}$ are all roots of (2.29) such that the $(n+1)$'th component of $\tilde{\boldsymbol{y}}(\lambda_i)$ is nonnegative and $\boldsymbol{B}^T\boldsymbol{B} + \lambda\boldsymbol{C}$ has exactly one negative and $n$ positive eigenvalues. We shall refer the

global solution of (2.27) to as the SRD-LS solution.

### 2.2.2 Improved Solution Using Iterative Re-weighting

#### The Algorithm

We now present a method for improved solutions over SRD-LS solutions. The method incorporates an iterative re-weighting procedure into the SRD-LS approach, hence it is in spirit similar to the IRWRS-LS approach described in Sec. 2.1.2. We begin by considering the weighted squared range-difference least squares (WSRD-LS) problem

$$\underset{\boldsymbol{x}\in R^n}{\text{minimize}} \sum_{i=1}^{m} w_i \left(-2\boldsymbol{a}_i^T\boldsymbol{x} - 2d_i\|\boldsymbol{x}\| - g_i\right)^2 \tag{2.30}$$

where weights $w_i$ for $i = 1, \ldots, m$ are fixed nonnegative constants. The counterpart of (2.27) for the problem (2.30) is given by

$$\underset{\boldsymbol{y}\in R^{n+1}}{\text{minimize}} \|\boldsymbol{B}_w\boldsymbol{y} - \boldsymbol{g}_w\| \tag{2.31a}$$

$$\text{subject to: } \boldsymbol{y}^T\boldsymbol{C}\boldsymbol{y} = 0 \tag{2.31b}$$

$$y_{n+1} \geq 0 \tag{2.31c}$$

where $\boldsymbol{g}_w = \boldsymbol{\Gamma}\boldsymbol{g}$, $\boldsymbol{B}_w = \boldsymbol{\Gamma}\boldsymbol{B}$, $\boldsymbol{\Gamma} = \text{diag}\{\sqrt{w_1}, \ldots, \sqrt{w_m}\}$, and $\boldsymbol{g}$, $\boldsymbol{B}$ are defined in (2.28). On comparing (2.31) with (2.27), it follows immediately that the global solver for problem (2.27) characterized by data set $\{\boldsymbol{B}, \boldsymbol{g}, \boldsymbol{C}\}$ can also be suited for solving problem (2.31) be used applying it to data set $\{\boldsymbol{B}_w, \boldsymbol{g}_w, \boldsymbol{C}\}$.

Concerning the assignment of weights $\{w_i, i = 1, \ldots, w_m\}$, we recall (2.23), (2.24) and observe that the $i$th term of the objective function in (2.30) can be written as

$$\begin{aligned}
&w_i \left(-2d_i\|\boldsymbol{x}\| - 2\boldsymbol{a}_i^T\boldsymbol{x} - g_i\right)^2 \\
=&w_i \left((d_i + \|\boldsymbol{x}\|)^2 - \|\boldsymbol{x} - \boldsymbol{a}_i\|^2\right)^2 \\
=&w_i \left(d_i + \|\boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{a}_i\|\right)\left(d_i + \|\boldsymbol{x}\| - \|\boldsymbol{x} - \boldsymbol{a}_i\|\right)
\end{aligned}$$

Clearly, the last expression above would become the $i$th term of the objective function in the RD-LS problem (2.22) if weights $w_i$ were set to

$$\frac{1}{\left(d_i + \|\boldsymbol{x}\| + \|\boldsymbol{x} - \boldsymbol{a}_i\|\right)^2}$$

so that the first two factors are cancelled out. This suggests that a realizable weight assignment for performing practically the same cancellation can be made by means of iterative re-weighting for problems (2.30) and (2.31) where the weights in the $k$th iteration are assigned to

$$w_i^{(k)} = \frac{1}{\left(d_i + \|\boldsymbol{x}_{k-1}\| + \|\boldsymbol{x}_{k-1} - \boldsymbol{a}_i\|\right)^2}, i = 1, \ldots, m \tag{2.32}$$

Based on the analysis above, a localization algorithm for range-difference measurements can be outlined as follows.

**Algorithm 2**

1) Input data: Sensor locations $\{\boldsymbol{a}_i, i = 0, 1, \ldots, m\}$ with $\boldsymbol{a}_0 = \boldsymbol{0}$, range-difference measurements $\{d_i, i = 1, \ldots, m\}$, maximum number of iterations $k_{max}$ and convergence tolerance $\xi$.

2) Generate data set $\{\boldsymbol{B}, \boldsymbol{g}, \boldsymbol{C}\}$ as

$$\boldsymbol{g} = \begin{pmatrix} d_1^2 - \|\boldsymbol{a}_1\|^2 \\ \vdots \\ d_m^2 - \|\boldsymbol{a}_m\|^2 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} -2\boldsymbol{a}_1^T & -2d_1 \\ \vdots & \vdots \\ -2\boldsymbol{a}_m^T & -2d_m \end{pmatrix}, \boldsymbol{C} = \begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{0}_{n\times 1} \\ \boldsymbol{0}_{1\times n} & -1 \end{pmatrix}.$$

Set $k = 1$, $w_i^{(1)} = 1$ for $i = 1, \ldots, m$.

3) Set $\boldsymbol{\Gamma}_k = \text{diag}\left(\sqrt{w_1^{(k)}}, \ldots, \sqrt{w_m^{(k)}}\right)$, $\boldsymbol{B}_w = \boldsymbol{\Gamma}_k \boldsymbol{B}$ and $\boldsymbol{g}_w = \boldsymbol{\Gamma}_k \boldsymbol{g}$.

4) Solve WSRD-LS problem

$$\begin{aligned} \underset{\boldsymbol{y}\in R^{n+1}}{\text{minimize}} \; & \|\boldsymbol{B}_w \boldsymbol{y} - \boldsymbol{g}_w\| \\ \text{subject to: } & \boldsymbol{y}^T \boldsymbol{C} \boldsymbol{y} = 0 \\ & y_{n+1} \geq 0 \end{aligned}$$

to obtain its global solution $\boldsymbol{x}_k$.

5) If $k = k_{max}$ or $\|\boldsymbol{x}_k - \boldsymbol{x}_{k-1}\| < \xi$, terminate and output $\boldsymbol{x}_k$ as the solution; otherwise, set $k = k + 1$, update weights $\{w_i^{(k)}, i = 1, \ldots, m\}$ as

$$w_i^{(k)} = \frac{1}{\left(d_i + \|\boldsymbol{x}_{k-1}\| + \|\boldsymbol{x}_{k-1} - \boldsymbol{a}_i\|\right)^2}$$

and repeat from Step 3).

It is evident that the complexity of the algorithm is practically equal to the complexity of the WSRD-LS solver involved in Step 4 times the number of iterations, $k$, which is typically in the range of 3 to 6. We shall call the solutions obtained from Algorithm 2 IRWSRD-LS solutions. Technical details with regard to solving problem (2.31) can be found in Appendix 1.

**A variant of Algorithm 2**

Like the case of range measurements, once the IRWSRD-LS solution is obtained by applying Algorithm 2, which is expected to be within a small vicinity of the true global solution of the RD-LS problem (2.22), the gap can be closed by running a good local method that takes the IRWSRD-LS solution as an initial point. Again, the Newton method is chosen for fast convergence, low complexity due to the extremely low dimension $n$, and the availability of closed-form formulas to compute the gradient and Hessian of $F(\boldsymbol{x})$ in (2.22).

Assuming $\boldsymbol{x} \neq \boldsymbol{a}_i$ for $i = 0, 1, \ldots, m$, the gradient and Hessian of $F(\boldsymbol{x})$ is found to be

$$\boldsymbol{g}(\boldsymbol{x}) = \sum_{i=1}^{m} c_i \left( \boldsymbol{q}_i - \tilde{(}\boldsymbol{x}) \right)$$

and

$$\boldsymbol{H}(\boldsymbol{x}) = \sum_{i=1}^{m} \left[ (\boldsymbol{q}_i - \tilde{(}\boldsymbol{x}))(\boldsymbol{q}_i - \tilde{(}\boldsymbol{x}))^T + c_i(\boldsymbol{Q}_{1i} + \boldsymbol{Q}_2) \right]$$

respectively, where

$$c_i = \|\boldsymbol{x} - \boldsymbol{a}_i\| - \|\boldsymbol{x}\|, \boldsymbol{q}_i = \frac{\boldsymbol{x} - \boldsymbol{a}_i}{\|\boldsymbol{x} - \boldsymbol{a}_i\|}, \tilde{\boldsymbol{x}} = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$$

and

$$\boldsymbol{Q}_{1i} = \frac{1}{\|\boldsymbol{x} - \boldsymbol{a}_i\|} \left( \boldsymbol{I} - \boldsymbol{q}_i \boldsymbol{q}^T \right), \boldsymbol{Q}_2 = \frac{1}{\|\boldsymbol{x}\|} \left( \boldsymbol{I} - \tilde{\boldsymbol{x}} \tilde{\boldsymbol{x}}^T \right)$$

To ensure the positive definiteness of Hessian, eigen-decomposition of $\boldsymbol{H}(\boldsymbol{x})$, namely,

$$\boldsymbol{H}(\boldsymbol{x}) = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^T$$

is performed, where $\boldsymbol{U}$ is orthogonal and $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \ldots, \lambda_n)$ with $\{\lambda_i, i = 1, \ldots, n\}$ being the eigenvalues of $\boldsymbol{H}(\boldsymbol{x})$. Let $\lambda_{min}$ be the smallest eigenvalue of $\boldsymbol{H}(\boldsymbol{x})$. If $\lambda_{min}$, then $\boldsymbol{H}(\boldsymbol{x})$ is positive definite and the Newton algorithm is carried out without

modification; if $\lambda_{min} \leq 0$, then the algorithm uses a slightly modified Hessian given by

$$\tilde{\boldsymbol{H}}(\boldsymbol{x}) = \boldsymbol{U}\tilde{\boldsymbol{\Lambda}}\boldsymbol{U}^T$$

where $\tilde{\boldsymbol{\Lambda}} = \text{diag}\left(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n\right)$ with

$$\tilde{\lambda}_i = \begin{cases} \lambda_i & \text{if } \lambda_i > 0 \\ \delta & \text{if } \lambda_i \leq 0 \quad i = 1, \ldots, m \end{cases}$$

and $\delta$ a small positive constant. Obviously, $\tilde{\boldsymbol{H}}(\boldsymbol{x})$ is guaranteed to be positive definite. In what follows, solutions obtained by the proposed two-step method are called *hybrid* IRWSRD-LS solutions.

### 2.2.3 Numerical Results

Performance of the proposed algorithms was evaluated and compared with the method of [15] by Monte Carlo simulations with a set-up similar to that of [15]. SRD-LS solutions were used as performance benchmarks for Algorithm 2 and its variant. In both cases the system consisted of $m = 11$ sensors $\{\boldsymbol{a}_i,\ i = 1, 2, \ldots, 10\}$ with $\boldsymbol{a}_0 = \boldsymbol{0}$ and other ten sensors placed randomly placed in the planar region $[-15; 15] \times [-15; 15]$, and a radiating source $\boldsymbol{x}_s$, located randomly in the region $[-10; 10] \times [-10; 10]$. Coordinates of the source and sensors were generated for each dimension following a uniform distribution. The range-difference measurements used to form matrix $\boldsymbol{B}$ in Step 2 of the Algorithm 2 were calculated as noise-contaminated range-difference measurements $d_i$ in (2.21). Measurement noise $\{\varepsilon_i,\ i = 1, \ldots, m\}$ was modelled as i.i.d. random variables with zero mean and variance $\sigma^2$, with $\sigma$ being one of five possible levels $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1\}$. Step 4 of Algorithm 2 was carried out using the SRD-LS algorithm in [15]. Accuracy of source location estimation was evaluated in terms of mean squared error in the form $\text{MSE} = E\{\|\boldsymbol{x}^* - \boldsymbol{x}_s\|\}$ where $\boldsymbol{x}_s$ denotes the exact source location and $\boldsymbol{x}^*$ is its estimation obtained by SRD-LS, IRWSRD-LS and *hybrid* IRWSRD-LS methods, respectively. Table 2.3 provides comparisons of these methods with SRD-LS, where each entry was averaged MSE over 1,000 Monte Carlo runs of the method. For the columns representing performance of the IRWSDR-LS and *hybrid* IRWSDR-LS methods each table entry lists their MSE and relative improvement over SRD-LS solutions in percentage, in the format of MSE (% Improvement). Again, the IRWSRD-LS solutions offer considerable improvement over

SRD-LS solutions. Further analysis of the data that was used to generate Table 2.3 illustrates the advantage of the IRWSR-LS (hybrid IRWSR-LS) solution over the SR-LS. Each entry in Table 2.4 is a standard deviation of the squared estimation errors aggregated over the same 1,000 Monte Carlo runs described above in Table 2.3 (where the MSE of the position estimation are shown). The results summarised in Table 2.4 suggest that, again, IRWSR-LS and hybrid IRWSR-LS outperform SR-LS.

Table 2.3: MSE of position estimation for SRD-LS, IRWSRD-LS and *hybrid* IRWSRD-LS methods

| $\sigma$ | SR - LS | IRWSR-LS (Im.,%) | *hybrid* IRWSR-LS (Im.,%) |
|---|---|---|---|
| 1e-04 | 1.38301598e-08 | 8.22705918e-09 (40) | 8.22705918e-09 (40) |
| 1e-03 | 1.60398717e-06 | 1.03880406e-06 (35) | 1.038804061e-06 (35) |
| 1e-02 | 1.11632818e-04 | 6.67785604e-05 (40) | 6.67265316e-05 (40) |
| 1e-01 | 1.20947651e-02 | 7.20891487e-03 (40) | 7.07178346e-03 (41) |
| 1e+0 | 1.57050323e+00 | 9.70756420e-01 (40) | 7.86289933e-01 (48) |

Table 2.4: Standard deviation of the squared estimation error for SRD-LS, IRWSRD-LS and *hybrid* IRWSRD-LS methods

| $\sigma$ | SR - LS | IRWSR-LS | *hybrid* IRWSR-LS |
|---|---|---|---|
| 1e-04 | 4.5624e-08 | 2.2446e-08 | 2.2446e-08 |
| 1e-03 | 3.9506e-06 | 3.1610e-06 | 3.1610e-06 |
| 1e-02 | 2.2710e-04 | 1.2812e-04 | 1.2812e-04 |
| 1e-01 | 3.0108e-02 | 1.8891e-02 | 1.8891e-02 |
| 1e+0 | 4.5781e+00 | 3.0597e+00 | 3.0597e+00 |

Figures 2.5 to 2.9 represent the histograms of the location estimation errors $\|\boldsymbol{x}^* - \boldsymbol{x}_s\|$ of the SR-LS solution (left images) and IRWSR-LS (right images) for all four noise levels with $\sigma$ being one of $\{10^{-3}, 10^{-2}, 10^{-1}, 1\}$, where $\boldsymbol{x}^*$ denotes the estimated location and $\boldsymbol{x}_s$ is the exact location of the source. Histograms that correspond to the results obtained by IRWSRD-LS are shifted closer towards 0 than those obtained by SR-LS, have smaller variance, and in most cases solutions obtained by IRWSRD-LS have fewer outliers.

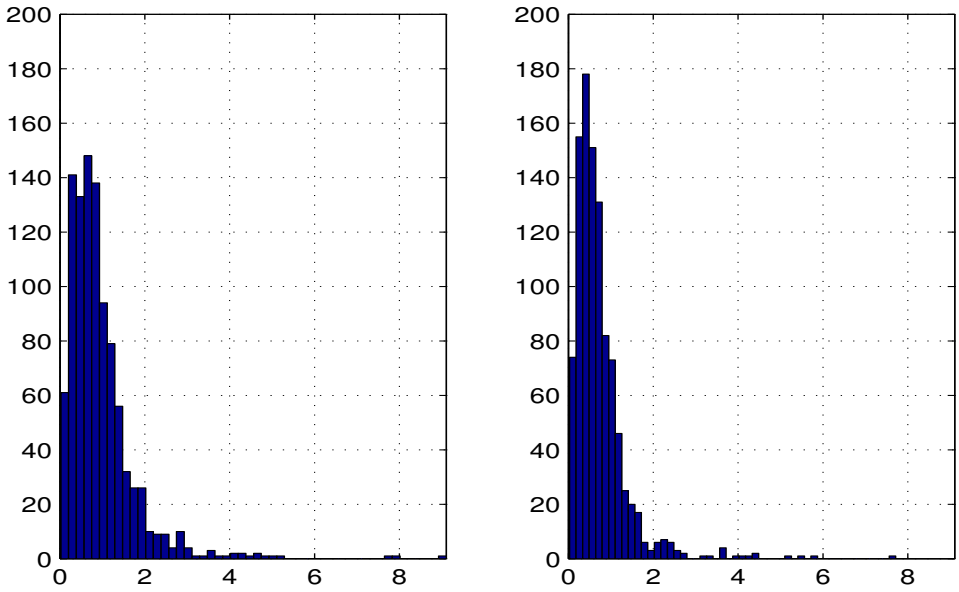



Figure 2.5: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 1$

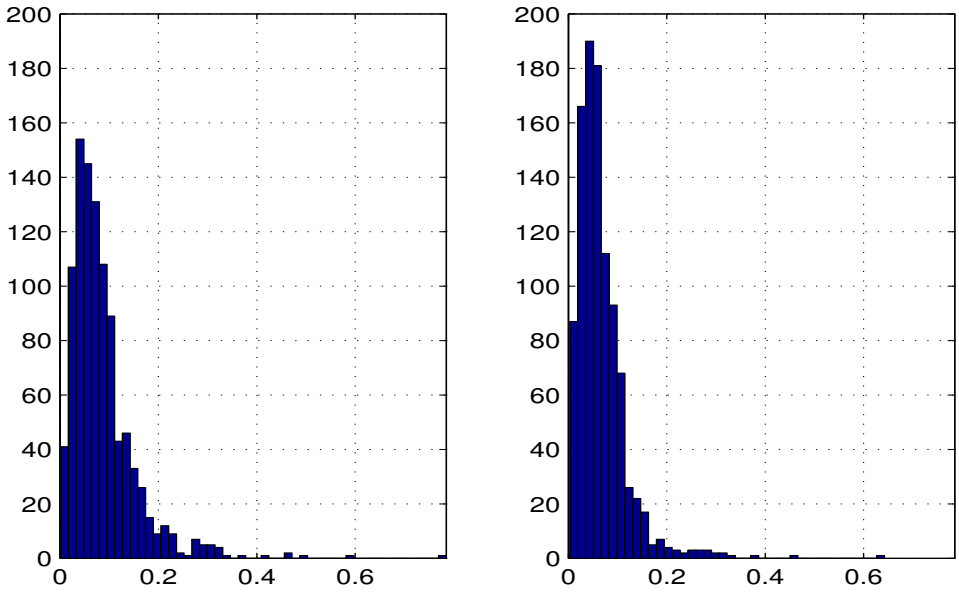



Figure 2.6: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 10^{-1}$

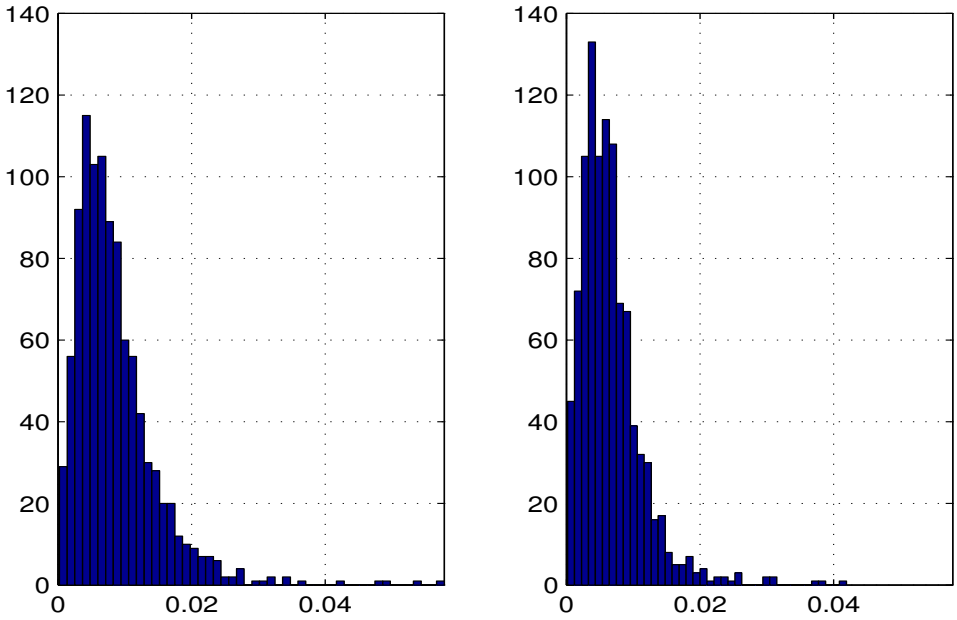



Figure 2.7: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 10^{-2}$

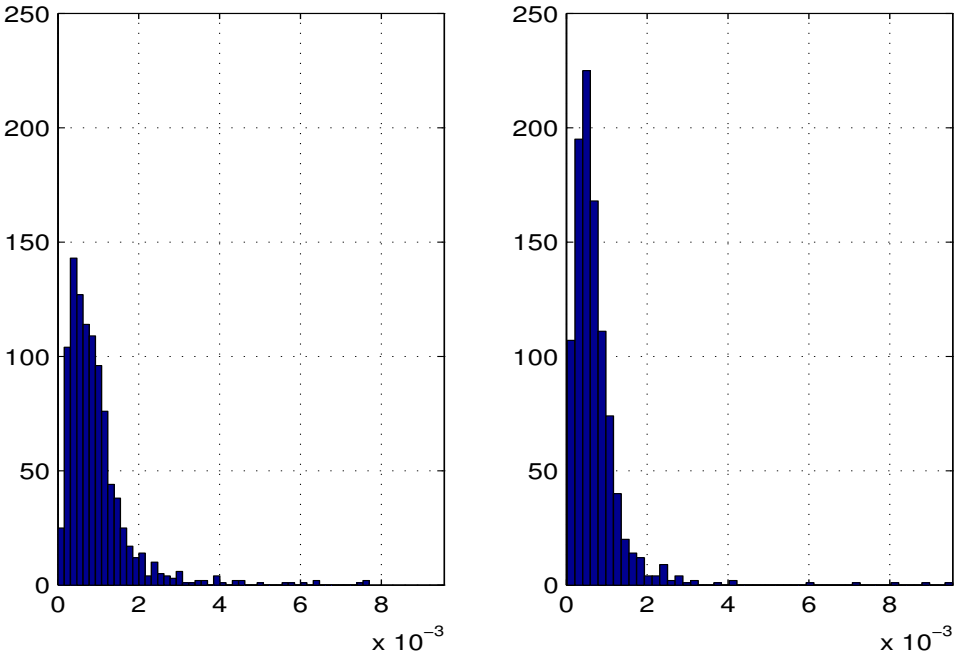



Figure 2.8: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 10^{-3}$

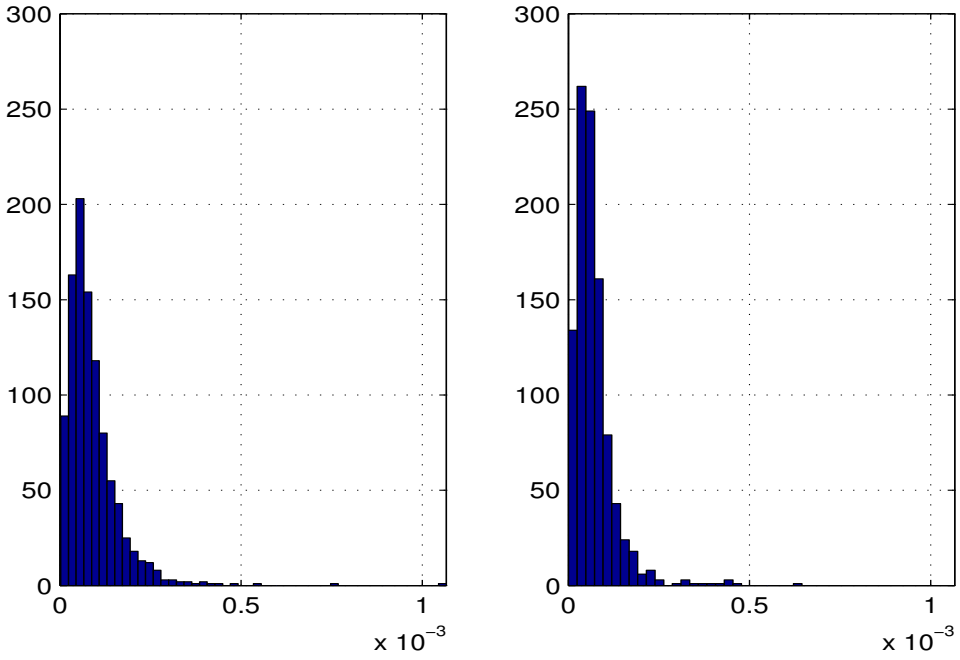



Figure 2.9: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, with standard deviation of noise $\sigma = 10^{-4}$

## 2.3 Extensions

Methods developed in this chapter for localization based on range measurements can also be adopted to solve the problem of single sourse localization using energy measurements [14]. Energy-based source localization, advocated in [7], [9], is motivated by the simple observation that the sound intensity received by a listener decreases when the distance between the sound source and the listener increases. By modeling the relation between sound intensity (energy) and distance from the sound source, one may estimate the source location using multiple energy readings at different known sensor locations. It is known that when the sound is propagating through the air, the acoustic energy emitted omnidirectionally from a sound source will attenuate at a rate that is inversely proportional to the square of the distance [7]. Using this fact and some simple manipulations, it is possible to obtain an equation in the vector of unknow source location $\boldsymbol{x}$ that is somewhat similar to (2.9). The rest of the section provides technical details of this reformulation.

### 2.3.1 Acoustic Energy Attenuation Model and Problem Statement

Let $m$ be a number of acoustic sensors involved. For consistency of notation, let $\boldsymbol{a}_i$ denote the known location of the sensor $i$ in space $R^n$ with $n = 2$ or 3. Each sensor measures the acoustic intensity radiated by a source $\boldsymbol{x} \in R^n$ over a time period $T = \frac{M}{f_s}$, where $M$ is the number of sample points used for estimating the acoustic enery and $f_s$ is the sampling frequency. Acoustic energy received by sensor $i$ over a time period $T$ can be represented as

$$r_i = g_i \frac{S}{\|\boldsymbol{x} - \boldsymbol{a}_i\|^\alpha} + \varepsilon_i \tag{2.33}$$

where $\|\boldsymbol{x} - \boldsymbol{a}_i\|$ is the Euclidean distance between the $i$th sensor and the source, and $g_i$ is a scaling factor that takes into account $i$th sensor gain. It is asssumed that the gain of individual sensors is either known, i.e. obtained at the sensor callibration stage, or is same for all sensors. In (2.33) $S$ is the unknown acoustic energy measured 1 unit distance away from the source, $\alpha$ is the energy decay factor and is usually assumed to have a value 2 [7], and $\varepsilon_i$ denotes the square of the background noise affecting the measurement of sensor $i$. If the number of sample points $M$ used for estimating the acoustic energy is large, the error $\varepsilon_i$ can be approximated well as a normal distribution with positive mean $\mu_i$ and variance $\sigma_i^2$ [7], [8]. More details on derivation and assumptions of this model can be found in [7] - [9], [29] and the references therein.

References [7, 8] argue that the maximum likelihood estimation of unknown parameters $\boldsymbol{\theta} = [\boldsymbol{x}^T S]^T$ can be obtained by solving the minimization problem

$$\underset{\theta}{\text{minimize}} \quad \ell(\boldsymbol{\theta}) = \|\boldsymbol{Z} - S\boldsymbol{H}\| \tag{2.34}$$

where

$$\boldsymbol{H} = \begin{pmatrix} \frac{g_1}{\sigma_1 \|\boldsymbol{x}-\boldsymbol{a}_1\|^2} \\ \frac{g_2}{\sigma_2 \|\boldsymbol{x}-\boldsymbol{a}_2\|^2} \\ \vdots \\ \frac{g_m}{\sigma_m \|\boldsymbol{x}-\boldsymbol{a}_m\|^2} \end{pmatrix}, \; \boldsymbol{Z} = \begin{pmatrix} \frac{y_1-\mu_1}{\sigma_1} \\ \frac{y_2-\mu_2}{\sigma_2} \\ \vdots \\ \frac{y_m-\mu_m}{\sigma_m} \end{pmatrix}$$

and $\boldsymbol{Z}$ are (estimated) normalized energy measurements for the case of the single radiating source. We remark that $\mu_i$ and $\sigma_i$ in (2.34) are assumed to be known.

### 2.3.2 Reformulation

There are two things to note about (2.34). First, (2.34) involves a *nonlinear* least square objective function because the vector $\boldsymbol{H}$ is a nonlinear function of the $n$ unknown source coordinates, where $n$ is the dimention of the location coordinates. Second, there are $m$ sensors reporting the acousting readings and there are a total of $n+1$ unknowns with $n+1 \leq m$, including the unknown acoustic energy $S$ radiated from the source. To eliminate the unknown source energy $S$ from formultion, reference [8] propose first to compute the ratio $k_{ij}$ of the calibrated energy readings from $i$th and $j$th sensor as

$$k_{ij} = \left(\frac{z_i/g_i}{z_j/g_j}\right)^{-1/\alpha} = \frac{\|\boldsymbol{x}-\boldsymbol{a}_i\|}{\|\boldsymbol{x}-\boldsymbol{a}_j\|} \tag{2.35}$$

for $i = 1, 2, \ldots, m-1$, and $j = i+1, \ldots, m$. For the case $0 < k_{ij} \neq 1$ all possible source coordinates $\boldsymbol{x}$ that form a solution to (2.35) reside on an $n$-dimensional hyper-sphere described by the equation:

$$\|\boldsymbol{x}-\boldsymbol{c}\|^2 = \rho_{ij}^2 \tag{2.36}$$

where the center $\boldsymbol{c}_{ij}$ and the radius $\rho_{ij}$ of the hyper-sphere associated with the sensors $i$ and $j$ are given by

$$\boldsymbol{c}_{ij} = \frac{\boldsymbol{a}_i - k_{ij}^2 \cdot \boldsymbol{a}_j}{1-k_{ij}^2},\ \rho_{ij} = \frac{k_{ij}\|\boldsymbol{a}_i - \boldsymbol{a}_j\|}{1-k_{ij}^2} \tag{2.37}$$

For the case when $k_{ij} \to 1$, the possible source locations $\boldsymbol{x}$ reside on the hyperplane between sensors $\boldsymbol{a}_i$ and $\boldsymbol{a}_j$, i.e.:

$$\boldsymbol{x}^T\boldsymbol{\gamma}_{ij} = \tau_{ij}$$

where $\boldsymbol{\gamma}_{ij} = \boldsymbol{a}_i - \boldsymbol{a}_j$ and $\tau_{ij} = (\|\boldsymbol{a}_i\|^2 - \|\boldsymbol{a}_j\|^2)/2$.

Let $I_1$ and $I_2$ be two index sets such that $0 < k_{ij} \neq 1$ for all $\{i,j\} \in I_1$ and $k_{ij} = 1$ for all $\{i,j\} \in I_2$ with $1 \leq i \leq m-1, i+1 \leq j \leq m$ and $I_1 \cap I_2 = \emptyset$. Let $L_1$ and $L_2$ denote the number of elements in sets $I_1$ and $I_2$ respectively (number of hyperspheres and hyperplanes) and $L_1 + L_2 = m(m-1)/2$. Then the unknown location of the source can be found via minimization of the following criterion which is equivalent to (2.34) [8]

$$\underset{\boldsymbol{x}}{\text{minimize}} \sum_{l_1=1}^{L_1} \left( \|\boldsymbol{x}-\boldsymbol{c}_{l_1}\|^2 - \rho_{l_1}^2 \right)^2 + \sum_{l_2=1}^{L_2} \left(\boldsymbol{x}^T\boldsymbol{\gamma}_{l_2} - \tau_{l_2}\right)^2 \tag{2.38}$$

For the brevity of notation, the double indexes $ij$ were replaced by single indexes $l_1$

and $l_2$. After some simple manupilations and necessary variable changes, (2.38) can be converted to the following constrained problem which is similar to (2.10)

$$\underset{\boldsymbol{y}\in R^{n+1}}{\text{minimize}} \, \|\boldsymbol{A}\boldsymbol{y}-\boldsymbol{b}\|^2 \tag{2.39a}$$

$$\text{subject to:} \quad \boldsymbol{y}^T\boldsymbol{D}\boldsymbol{y}+2\boldsymbol{f}^T\boldsymbol{y}=0 \tag{2.39b}$$

where

$$\boldsymbol{y}=\begin{pmatrix} \boldsymbol{x} \\ \|\boldsymbol{x}\|^2 \end{pmatrix}, \; \boldsymbol{A}=\begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix}, \; \boldsymbol{b}=\begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix} \tag{2.40}$$

$$\boldsymbol{D}=\begin{pmatrix} \boldsymbol{I}_{n\times n} & \boldsymbol{0}_{n\times 1} \\ \boldsymbol{0}_{1\times n} & 0 \end{pmatrix}, \; \boldsymbol{f}=\begin{pmatrix} \boldsymbol{0} \\ -0.5 \end{pmatrix}$$

and submatrices of $\boldsymbol{A}$ and elements of $\boldsymbol{b}$ are formed as follows:

$$\boldsymbol{A}_1=\begin{pmatrix} -2\boldsymbol{c}_1^T & 1 \\ \vdots & \vdots \\ -2\boldsymbol{a}_{L_1}^T & 1 \end{pmatrix}, \; \boldsymbol{b}_1=\begin{pmatrix} \rho_1^2-\|\boldsymbol{c}_1\|^2 \\ \vdots \\ \rho_{L_1}^2-\|\boldsymbol{c}_{L_1}\|^2 \end{pmatrix} \tag{2.41}$$

$$\boldsymbol{A}_2=\begin{pmatrix} \gamma_1^T & 0 \\ \vdots & \vdots \\ \gamma_{L_2}^T & 0 \end{pmatrix}, \; \boldsymbol{b}_2=\begin{pmatrix} \tau_1 \\ \vdots \\ \tau_{L_2} \end{pmatrix}$$

The unconstrained problem in (2.38), due to mathematical analogy with (2.9), can be solved using Algorithm 1 developed in Section 3.1.

# Chapter 3

# Penalty Convex-Concave Procedure for Source Localization

In connection with the problem of localizing a single radiating source based on range measurements, in this chapter we explore special structure of the cost function of an unconstrained least squares (LS) formulation and show that it is well suited in a setting known as difference-of-convex-functions (DC) programming. In the literature, the DC programming is sometimes referred to as convex-concave procedure. Our focus in this chapter will be placed on the localization problem based on range measurements. We present an algorithm for solving the LS problem at hand based on a penalty convex-concave procedure (PCCP) [33] that accommodates infeasible initial points in solving a fairly large class of *nonconvex* constrained problems. Algorithmic details are provided to show that the PCCP-based formulation is tailored to the localization problem at hand. These include additional constraints that enforce the algorithms iteration path towards the LS solution, and several strategies to secure good initial points. Numerical results are presented to demonstrate that the proposed algorithm offers substantial performance improvement relative to some best known results from the literature.

## 3.1 Problem Statement and Review of Related Work

Typically non-survey based localization techniques compute the location estimates in two steps: range/angle estimation and tri-lateration/angulation [44]. In general, the

range estimates can be based on different types of measurements, e.g. received signal strength (RSS), or time of arrival (TOA). This chapter will focus on the problem of range-based localization given the TOA information. In the TOA method, the one-way propagation time of the signal traveling between radiating source and the sensor node is measured. Each TOA measurement then provides a circle centered at the sensor node on which the source of the signal must lie. With three or more sensor nodes the measurements are converted into a set of circular equations that, with knowledge of the geometry of the sensor network, allow to determine the unknown source position [44]. The accuracy of the positioning depends on the quality of the range measurements, network geometry, and the performace of the localization algorithm. In real-world situations, multipath and non line-of-sight (NLOS) propagation are two major sources of error, which can introduce large biases in the TOA measurements and result in unreliable position esitmation [24]. In fact, mitigation of the impairments due to multipath and/or NLOS is another key research topic in wireless location and recent works in this area have reported some promising results [38]. Ultra-wideband (UWB) technology has the potential to deliver very accurate range measurements, thus enabling accurate positioning [37, 38, 39]. As a result, we assume that the multipath and NLOS errors in the TOA measurements have been successfully mitigated.

The source localization problem discussed in this chapter involves a given array of $m$ sensors placed in the $n = 2$ or 3 dimentional space with coordinates specified by $\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m, \boldsymbol{a}_i \in R^n\}$. Each sensor measures its distance to a radiating source $\boldsymbol{x} \in R^n$. Throughout it is assumed that only noisy copies of the distance data are available, hence the *range measurements* obey the model

$$r_i = \|\boldsymbol{x} - \boldsymbol{a}_i\| + \varepsilon_i, \quad i = 1, \ldots, m. \tag{3.1}$$

where $\varepsilon_i$ denotes the unknown noise that has occurred when the $i$th sensor measures its distance to source $\boldsymbol{x}$. Let $\boldsymbol{r} = [r_1\ r_2 \ldots r_m]^T$ and $\boldsymbol{\varepsilon} = [\varepsilon_1\ \varepsilon_2 \ldots \varepsilon_m]^T$. The source localization problem can be stated as to estimate the exact source location $\boldsymbol{x}$ from noisy range measurements $\boldsymbol{r}$.

The nonlinear least squares (NLLS) estimate refers to the solution of the problem

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad F(\boldsymbol{x}) = \sum_{i=1}^{m} (r_i - \|\boldsymbol{x} - \boldsymbol{a}_i\|)^2 \tag{3.2}$$

If ranging errors $\varepsilon_i$ are i.i.d. random variables that follow Gaussian distribution with zero mean and covariance matrix proportional to the identity matrix, then the NLLS estimate becomes identical to the maximum likelihood estimate. The NLLS formulation is also geometrically meaningful and has been often used as a benchmark to compare new algorithms [15, 37].

In Sec.2.1 of the thesis, it is demonstrated that these problems are hard to solve globally. Many relaxation and approximation methods were developed that offer either lower computational complexity, robustness against positive bias in the distance estimates due to non line-of-site situations, or better performance compared to standard unconstrain optimization methods applied to the NLLS problem. Convex relaxation of a nonconvex problem in (3.2) to an SDP problem and solution methods for *squared* range LS problems are discussed in detail in [47] and Sec.2.1 and of the thesis. Another localization approach that has received considerable interest applies classical multidimensional scaling (MDS) algorithm or its variants to the problem at hand [24, 25, 26, 27].

Multidimensional scaling is a field of study concerning the search of points in a low dimensional space that represent the objects of interest and the pairwise distances between the points (objects) (as measure of dissimilarities) match a set of given values. As such MDS has been an attractive technique for analyzing experimental data in physical, biological, and behavioral science [24]. The classical MDS is a subset of MDS techniques where the relative coordinates of points are determined given only their pairwise Euclidean distances.

When applied to the localization problem at hand, classical MDS starts with constructing a multidimentional similarity matrix. Let $\boldsymbol{X}$ denote an $m$ x $n$ distance matrix

$$\boldsymbol{X} = \begin{pmatrix} (\boldsymbol{x} - \boldsymbol{a}_1)^T \\ (\boldsymbol{x} - \boldsymbol{a}_2)^T \\ \vdots \\ (\boldsymbol{x} - \boldsymbol{a}_m)^T \end{pmatrix}$$

The multidimentional similarity matrix is then defined by $\boldsymbol{D} = \boldsymbol{X}\boldsymbol{X}^T$ which can also be expressed in terms of pairwise distances between sensor nodes and error-free range measurements. Since the exact distances between the sourse $\boldsymbol{x}$ and sensors are not available, the noisy range measurements $r_i$ are used to construct an approximation

of $\boldsymbol{D}$ as

$$\hat{\boldsymbol{D}} = \frac{1}{2}\begin{pmatrix} 2r_1^2 & r_1^2 + r_2^2 - \|\boldsymbol{a}_1 - \boldsymbol{a}_2\|^2 & \dots & r_1^2 + r_m^2 - \|\boldsymbol{a}_1 - \boldsymbol{a}_m\|^2 \\ r_1^2 + r_2^2 - \|\boldsymbol{a}_1 - \boldsymbol{a}_2\|^2 & 2r_2^2 & \dots & r_2^2 + r_m^2 - \|\boldsymbol{a}_2 - \boldsymbol{a}_m\|^2 \\ \vdots & \vdots & \ddots & \vdots \\ r_1^2 + r_m^2 - \|\boldsymbol{a}_1 - \boldsymbol{a}_m\|^2 & r_2^2 + r_m^2 - \|\boldsymbol{a}_1 - \boldsymbol{a}_m\|^2 & \dots & 2r_m^2 \end{pmatrix}$$

Because matrix $\hat{\boldsymbol{D}}$ is symmetric, it admits the orthogonal eigen-decomposition

$$\hat{\boldsymbol{D}} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{U}^T$$

where $\boldsymbol{\Lambda} = \text{diag}\,(\lambda_1, \lambda_2, \dots, \lambda_m)$ is the diagonal matrix of eigenvalues of $\hat{\boldsymbol{D}}$ with $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$, and $\boldsymbol{U} = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2 \dots \boldsymbol{u}_m]$ is an orthonormatl matrix whose columns are the corresponding eigenvectors. Since the rank of the ideal $\boldsymbol{D}$ is 2, an LS estimate of $\boldsymbol{X}$, denoted by $\boldsymbol{X}_r$, can be computed up to an arbitrary rotation as a solution to the following problem [24]

$$\boldsymbol{X}_r = \arg\min_{\tilde{\boldsymbol{X}}} \|\hat{\boldsymbol{D}} - \tilde{\boldsymbol{X}}\tilde{\boldsymbol{X}}^T\|_F^2 = \boldsymbol{U}_s\boldsymbol{\Lambda}_s^{(1/2)}$$

where $\tilde{\boldsymbol{X}}$ is the variable matrix for $\boldsymbol{X}$, $\|.\|_F$ represents the Frobenius norm, $\boldsymbol{U}_s = [\boldsymbol{u}_1 \quad \boldsymbol{u}_2]$ corresponds to the signal subspase, and $\boldsymbol{\Lambda}^{(1/2)} = \text{diag}(\lambda_1^{(1/2)}, \lambda_2^{(1/2)})$. In practical situations of nonzero range errors, the relationship between $\boldsymbol{X}_r$ and $\boldsymbol{X}$ is then

$$\boldsymbol{X} \approx \boldsymbol{X}_r\boldsymbol{\Omega}$$

where $\boldsymbol{\Omega}$ is an unknown rotation matrix to be determined. The estimate of the unknown rotation matrix $\boldsymbol{\Omega}$ and source location $\boldsymbol{x}$ can be obtained by solving an overdetermined system of linear equations [24]. In the absence of noise the symmetric $\hat{\boldsymbol{D}}$ is identical to $\boldsymbol{D}$, is positive semi-definite, and has a rank of 2. In the practical situations of nonzero range errors, $\hat{\boldsymbol{D}}$ will have a full rank.

Other methods based on MDS include a generalized subspace approach by So and Chan [27], that performs position estimation based on the noise subspace. A subspase-based weighting Lagrangian multiplier estimator [25] reduces computational complexity by avoiding the process of eigendecomposition or inverse computation, but it requires some a priori knowledge about noise statistic to construct the weighting matrix. On the other hand, the distributed weighted-multidimentional scaling

(DW-MDS) [26] adds a penalty term to the standard MDS objective function which accounts for prior knowledge about node locations. Although these methods work well in general and can be efficient in terms of complexity, they are found to produce poor performance in certain sensor deployments [37].

In this chapter, we focus on the least squares formulation for the localization problem, where the $l_2$-norm of the residual errors is minimized in a setting known as difference-of-convex-functions programming. The problem at hand is then solved by applying a penalty convex-concave procedure (PCCP) in a succesive manner [48].

## 3.2 Fitting the Localization Problem into a CCP Framework

### 3.2.1 Basic Convex-Concave Procedure

The CCP refers to an effective heuristic method to deal with a class of difference of convex (DC) programming problems, which assume the form

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}) - g(\boldsymbol{x}) && \text{(3.3a)} \\ \text{subject to:} \quad & f_i(\boldsymbol{x}) \leq g_i(\boldsymbol{x}) \quad \text{for: } i = 1, 2, \ldots, m && \text{(3.3b)} \end{aligned}$$

where $f(\boldsymbol{x}), g(\boldsymbol{x}), f_i(\boldsymbol{x}), g_i(\boldsymbol{x})$ for $i = 1, 2, \ldots, m$ are convex. A DC program is not convex unless the functions $g$ and $g_i$ are affine, and therefore is hard to solve in general. The class of DC functions is very broad. For example, any $C^2$ function can be expressed as a difference of convex functions [41]. Classes of problems that can be expressed as difference of convex programming among others include Boolean linear program, circle packing, circuit layout, and multi-matrix proncipal component analysis, examples of which can be found in [33] and references therein.

The basic CCP algorithm is an iterative procedure including two key steps (in the $k$-th iteration where iterate $\boldsymbol{x}_k$ is known):

(i) Convexification of the objective function and constraints by replacing $g(\boldsymbol{x})$ and $g_i(\boldsymbol{x})$, respectively, with their affine approximations

$$\hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) = g(\boldsymbol{x}_k) + \bigtriangledown g(\boldsymbol{x}_k)^T (\boldsymbol{x} - \boldsymbol{x}_k) \qquad \text{(3.4a)}$$

and

$$\hat{g}_i(\boldsymbol{x}, \boldsymbol{x}_k) = g_i(\boldsymbol{x}_k) + \bigtriangledown g_i(\boldsymbol{x}_k)^T(\boldsymbol{x} - \boldsymbol{x}_k) \quad \text{for: } i = 1, 2, \ldots, m \tag{3.4b}$$

(ii) Solving the convex problem

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}) - \hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) && \text{(3.5a)} \\ \text{subject to:} \quad & f_i(\boldsymbol{x}) - \hat{g}_i(\boldsymbol{x}, \boldsymbol{x}_k) \leq 0 && \text{(3.5b)} \\ & \text{for: } i = 1, 2, \ldots, m \end{aligned}$$

Because of the convexity of all the functions involved, it can be shown that the basic CCP is a descent algorithm and the iterates $\boldsymbol{x}_k$ converge to the critical point of the original problem [33]. In fact, the global convergence analysis for CCP has also been studied [34, 35]. Note that functions $g(\boldsymbol{x})$ and $g_i(\boldsymbol{x})$ for $i = 1, 2, \ldots, m$ are required to be convex but not necessarily differentiable. If any of $g(\boldsymbol{x})$ or $g_i(\boldsymbol{x})$ are not differentiable at some point $\tilde{\boldsymbol{x}}$ then the corresponding term $\bigtriangledown g(\tilde{\boldsymbol{x}})$ (or $\bigtriangledown g_i(\tilde{\boldsymbol{x}})$) is replaced by a subgradient of $g(\boldsymbol{x})$ (or $g_i(\boldsymbol{x})$) at point $\tilde{\boldsymbol{x}}$.

Let $\mathcal{D}$ be a nonempty set in $R^n$. A vector $\boldsymbol{h} \in R^n$ is said to be a subgradient of a convex function $f : \mathcal{D} \to R$ at $\boldsymbol{x} \in \mathcal{D}$ if

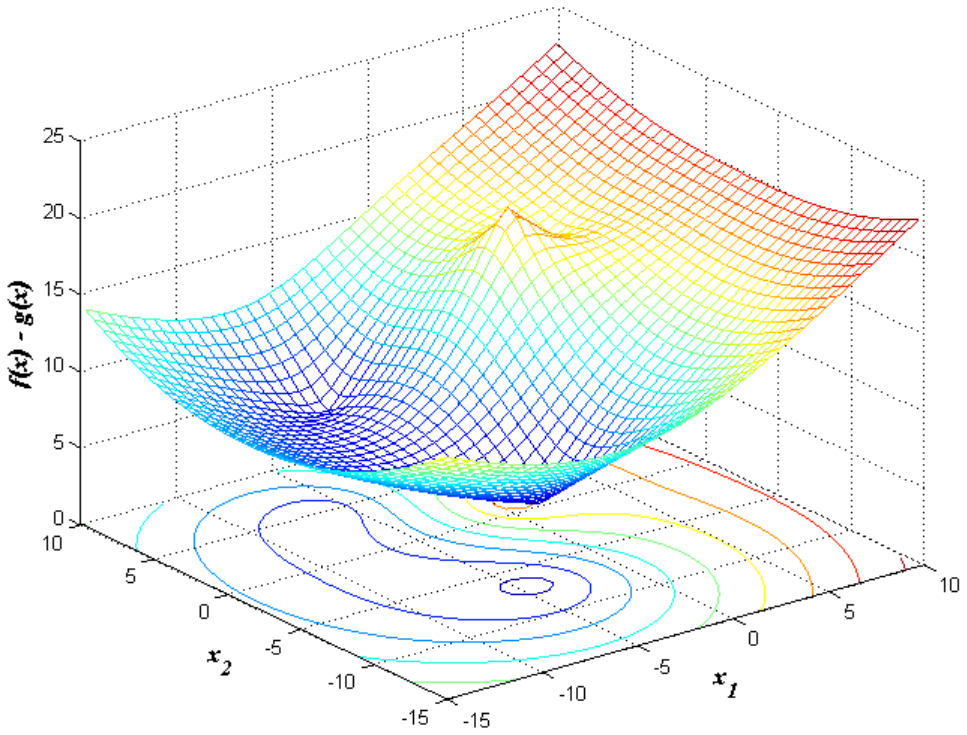
$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \boldsymbol{h}^T(\boldsymbol{y} - \boldsymbol{x}) \text{ for all } \boldsymbol{y} \in \mathcal{D}$$

Geometrically, the subgradients at a point $\boldsymbol{x}$ for the case where the convex function $f(x)$ is not differentiable correspond to different tangent lines at $\boldsymbol{x}$ [42].

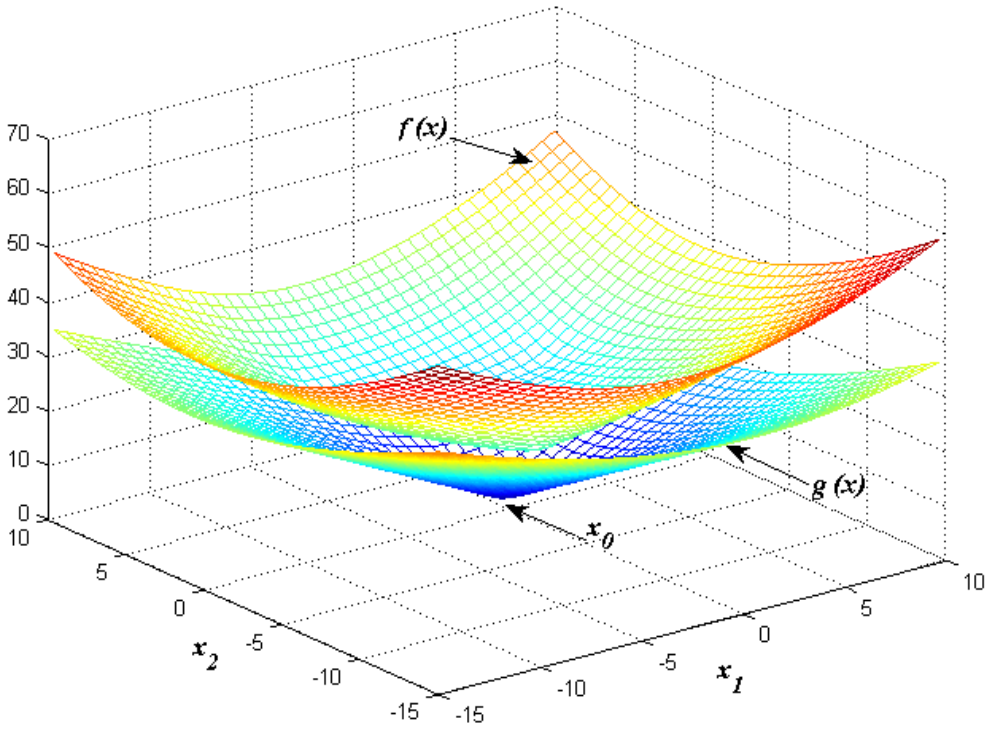
Figure 3.1 shows an example of the CCP approach for an unconstrained DC problem in the form of (3.3a), where $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ are given by:

$$\begin{aligned} f(\boldsymbol{x}) &= \|\boldsymbol{x} - \boldsymbol{a}_1\| + \|\boldsymbol{x} - \boldsymbol{a}_2\| + \|\boldsymbol{x} - \boldsymbol{a}_3\| \\ g(\boldsymbol{x}) &= \|2\boldsymbol{x} - \boldsymbol{a}_4\| \end{aligned}$$

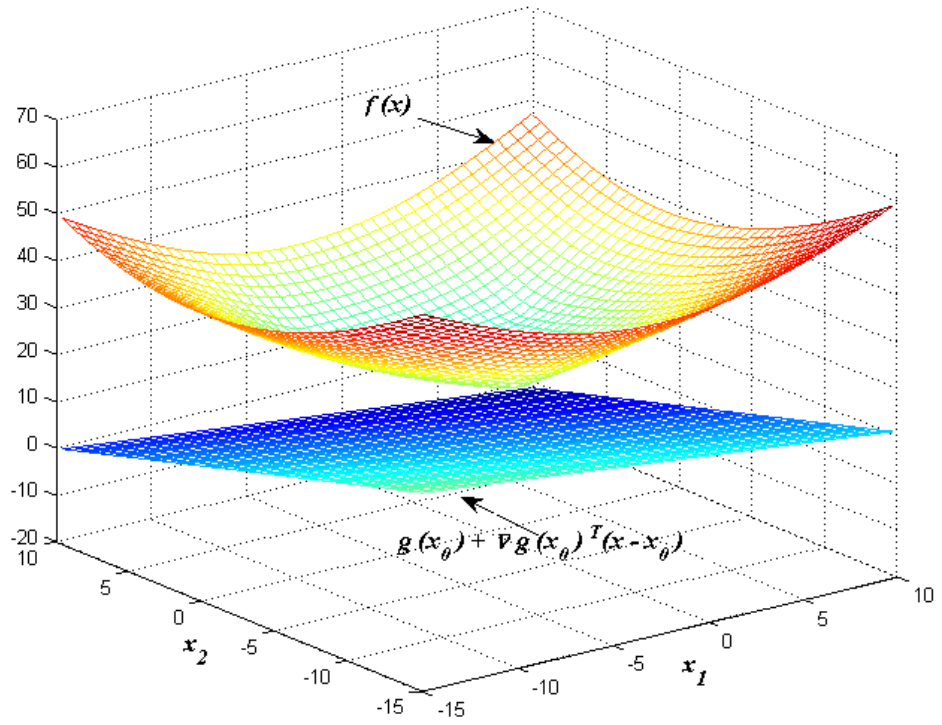
with $\boldsymbol{a}_1 = [3\ 2]^T$, $\boldsymbol{a}_2 = [6\ 5]^T$, $\boldsymbol{a}_3 = [4\ 7]^T$, and $\boldsymbol{a}_4 = [1\ 2]^T$. In this figure, the original nonconvex problem is transferred to a convex problem by replacing a nonconvex part $(-g(\boldsymbol{x}))$ by its affine approximation around the point $\boldsymbol{x}_0 = [0\ 0]^T$.



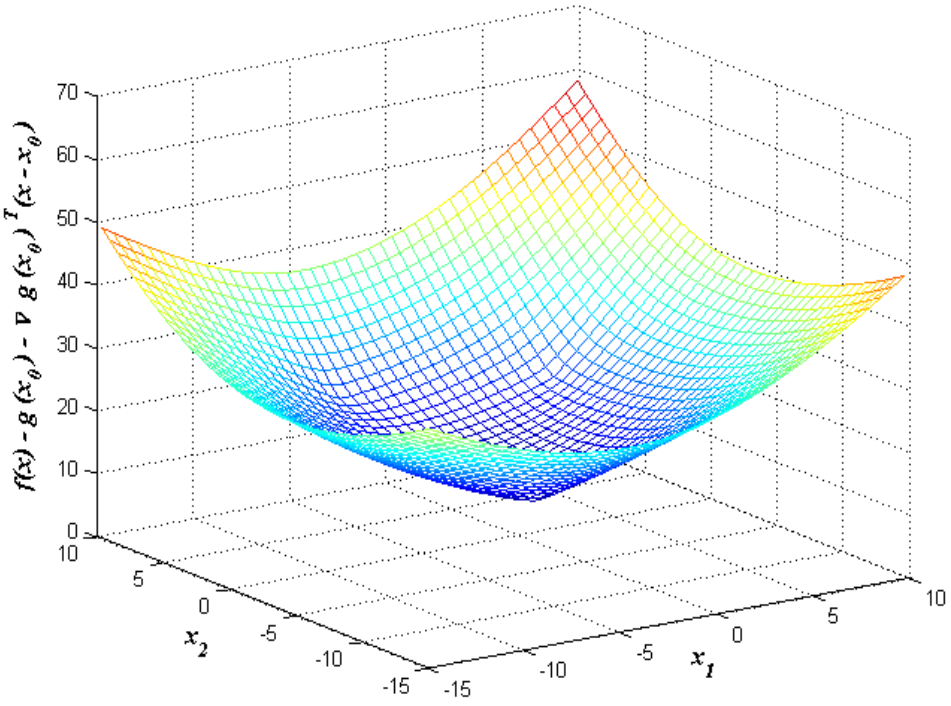

(a) A nonconvex function in the form of the difference of two convex functions and its contour plot.




(b) Separation of the nonconvex function into two convex functions $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$.




(c) First order approximation of $g(\boldsymbol{x})$.




(d) A convex approximation of the original nonconvex function at $\boldsymbol{x}_0 = [0\ 0]^T$.

Figure 3.1: An example of the CCP procedure (re-generated based on [36]).

The basic CCP requires a *feasible* initial point $\boldsymbol{x}_0$ (in the sense that $\boldsymbol{x}_0$ satisfies (3.5b) for $i = 1, 2, \ldots, m$) to start the procedure. By introducing additional slack variables, a penalty CCP has been adopted to accept infeasible initial points. In what follows, we reformulate our localization problem to fit it into the basic CCP framework. Bounds on squared measurement errors as well as penalty terms are then imposed, and a PCCP-based algorithm is developed for solving the problem.

### 3.2.2 Problem Reformulation

We begin by re-writing the NLLS objective function in (3.2) up to a constant as

$$F(\boldsymbol{x}) = m\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\sum_{i=1}^{m}\boldsymbol{a}_i - 2\sum_{i=1}^{m} r_i\|\boldsymbol{x} - \boldsymbol{a}_i\| \tag{3.6}$$

The objective in (3.6) is not convex. This is because, for points $\boldsymbol{x}$ that are not coincided with $\boldsymbol{a}_i$ for $1 \leq i \leq m$ , the Hessian of $F(\boldsymbol{x})$ is given by

$$\bigtriangledown^2 F(\boldsymbol{x}) = 2m\boldsymbol{I} + 2\sum_{i=1}^{m}\frac{r_i}{\|\boldsymbol{x} - \boldsymbol{a}_i\|^3}\cdot\Big((\boldsymbol{x} - \boldsymbol{a}_i)(\boldsymbol{x} - \boldsymbol{a}_i)^T - \|\boldsymbol{x} - \boldsymbol{a}_i\|^2\boldsymbol{I}\Big)$$

which is obviously not always positive semidefinite. On the other hand, by defining

$$\begin{aligned} f(\boldsymbol{x}) &= m\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\sum_{i=1}^{m}\boldsymbol{a}_i \\ g(\boldsymbol{x}) &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x} - \boldsymbol{a}_i\| \end{aligned} \tag{3.7}$$

the objective in (3.6) can be expressed as

$$F(\boldsymbol{x}) = f(\boldsymbol{x}) - g(\boldsymbol{x})$$

with both $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ convex, hence it fits naturally into (3.3). Note that $g(\boldsymbol{x})$ in (3.7) is not differentiable at the point where $\boldsymbol{x} = \boldsymbol{a}_i$ for some $1 \leq i \leq m$, thus we replace the term $\bigtriangledown g(\boldsymbol{x}_k)$ in (3.4b) by a subgradient [43] of $g(\boldsymbol{x})$ at $\boldsymbol{x}_k$, denoted by $\partial g(\boldsymbol{x}_k)$, as

$$\partial g(\boldsymbol{x}_k) = 2\sum_{i=1}^{m} r_i\partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\|$$

where

$$\partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| = \begin{cases} \dfrac{\boldsymbol{x}_k - \boldsymbol{a}_i}{\|\boldsymbol{x}_k - \boldsymbol{a}_i\|}, & \text{if } \boldsymbol{x}_k \neq \boldsymbol{a}_i \\ \mathbf{0}, & \text{otherwise} \end{cases}$$

Hence $\hat{g}(\boldsymbol{x}, \boldsymbol{x}_k)$ in (3.4b) is given by

$$\begin{aligned} \hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}_k - \boldsymbol{a}_i\| + 2\left(\boldsymbol{x} - \boldsymbol{x}_k\right)^T \sum_{i=1}^{m} r_i \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| \\ &= 2\boldsymbol{x}^T \sum_{i=1}^{m} r_i \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| + c \end{aligned}$$

where $c$ is a constant given by

$$\begin{aligned} c &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - 2\boldsymbol{x}_k^T \sum_{i=1}^{m} r_i \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| \\ &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - 2\sum_{i=1}^{m} r_i \boldsymbol{x}_k^T \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| \\ &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - 2\sum_{i=1}^{m} r_i \left(\boldsymbol{x}_k^T - \boldsymbol{a}_i + \boldsymbol{a}_i\right)^T \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| \\ &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - 2\sum_{i=1}^{m} r_i \left(\boldsymbol{x}_k^T - \boldsymbol{a}_i\right)^T \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - 2\sum_{i=1}^{m} r_i \boldsymbol{a}_i^T \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| \\ &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - 2\sum_{i=1}^{m} r_i \boldsymbol{a}_i^T \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| \\ &= -2\sum_{i=1}^{m} r_i \boldsymbol{a}_i^T \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\|. \end{aligned}$$

The convex approximation of the objective in (3.6) can now be derived as

$$\begin{aligned} \hat{F}(\boldsymbol{x}) &= f(\boldsymbol{x}) - \hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) \\ &= m\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T \sum_{i=1}^{m} \boldsymbol{a}_i - 2\boldsymbol{x}^T \sum_{i=1}^{m} r_i \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\| + c \end{aligned}$$

It follows that, up to a multiplicative factor $1/m$ and an additive constant term, the convex objective function in (3.5b) can be written as

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \hat{F}(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{v}_k \tag{3.8}$$

where

$$\boldsymbol{v}_k = \bar{\boldsymbol{a}} + \frac{1}{m}\sum_{i=1}^{m} r_i \partial \|\boldsymbol{x}_k - \boldsymbol{a}_i\|, \quad \bar{\boldsymbol{a}} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{a}_i \tag{3.9}$$

It is rather straightforward to see that given $\boldsymbol{x}_k$ (in the $k$-th iteration) the solution of the quadratic problem (3.8) can be obtained as

$$\boldsymbol{x}_{k+1} = \bar{\boldsymbol{a}} + \frac{1}{m}\sum_{i=1}^{m} r_i \partial \|\boldsymbol{x}_k - \boldsymbol{a}_i\|. \tag{3.10}$$

### 3.2.3 Imposing Error Bounds and Penalty Terms

The algorithm being developed can be enhanced by imposing a bound on each squared measurement error, namely

$$\left(\|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i\right)^2 \leq \delta_i^2 \tag{3.11}$$

which leads to

$$\|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i - \delta_i \leq 0 \tag{3.12a}$$

$$r_i - \delta_i \leq \|\boldsymbol{x} - \boldsymbol{a}_i\| \tag{3.12b}$$

for $1 \leq i \leq m$. Placing such bounds means that as iterations proceed the new iterates (coordinates of possible source locations) are restricted to lie within a physically meaningfull region determined by parameters $\delta_i$.

Note that the constraints in (3.12a) are convex and fit into the form of basic CCP in (3.5b) with $f_i(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i - \delta_i$ and $g_i(\boldsymbol{x}) = 0$, while those in (3.12b) are in the form of (3.3) with $f_i(\boldsymbol{x}) = r_i - \delta_i$ and $g_i(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}_i\|$. Following CCP (see (3.4b)), $g_i(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}_i\|$ is linearized around iterate $\boldsymbol{x}_k$ to

$$\hat{g}_i(\boldsymbol{x}, \boldsymbol{x}_k) = \|\boldsymbol{x}_k - \boldsymbol{a}_i\| + \partial \|\boldsymbol{x}_k - \boldsymbol{a}_i\|^T (\boldsymbol{x} - \boldsymbol{x}_k)$$

and (3.12b) is convexified as

$$r_i - \delta_i \leq \|\boldsymbol{x}_k - \boldsymbol{a}_i\| + \partial \|\boldsymbol{x}_k - \boldsymbol{a}_i\|^T (\boldsymbol{x} - \boldsymbol{x}_k)$$

which now fits into (3.5b), or equivalently

$$-\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\|^T (\boldsymbol{x} - \boldsymbol{x}_k) + r_i - \delta_i \leq 0 \tag{3.13}$$

We remark that constraint (3.13) is not only convex but also tighter than (3.12b). As a matter of fact, the convexity of the norm $\|\boldsymbol{x} - \boldsymbol{a}_i\|$ implies that it obeys the property

$$\|\boldsymbol{x} - \boldsymbol{a}_i\| \geq \|\boldsymbol{x}_k - \boldsymbol{a}_i\| + \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\|^T (\boldsymbol{x} - \boldsymbol{x}_k)$$

Therefore, a point $\boldsymbol{x}$ satisfying (3.13) automatically satisfies (3.12b). Summarizing, the convexified problem in the $k$-th iteration can be stated as

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{v}_k && \text{(3.14a)} \\ \text{subject to:} \quad & \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i - \delta_i \leq 0 && \text{(3.14b)} \end{aligned}$$

$$-\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\|^T (\boldsymbol{x} - \boldsymbol{x}_k) + r_i - \delta_i \leq 0 \tag{3.14c}$$

A technical problem making the formulation in (3.14) difficult to implement is that it requires a feasible initial point $\boldsymbol{x}_0$. The problem can be overcome by introducing nonnegative slack variables $s_i \geq 0, \hat{s}_i \geq 0$, for $i = 1, \ldots, m$ into the constraints in (3.14b) and (3.14c) to replace their right-hand sides (which are zeros) by relaxed upper bounds (as these new bounds themselves are nonnegative variables). This leads to a *penalty* CCP (PCCP) based formulation as follows:

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{s},\hat{\boldsymbol{s}}}{\text{minimize}} \quad & \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{v}_k + \tau_k \sum_{i=1}^{m} (s_i + \hat{s}_i) && \text{(3.15a)} \\ \text{subject to:} \quad & \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i - \delta_i \leq s_i && \text{(3.15b)} \end{aligned}$$

$$-\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - \frac{(\boldsymbol{x}_k - \boldsymbol{a}_i)^T}{\|\boldsymbol{x}_k - \boldsymbol{a}_i\|} (\boldsymbol{x} - \boldsymbol{x}_k) + r_i - \delta_i \leq \hat{s}_i \tag{3.15c}$$

$$s_i \geq 0, \hat{s}_i \geq 0, \text{ for: } i = 1, 2, \ldots, m \tag{3.15d}$$

where the weight $\tau_k \geq 0$ increases as iterations proceed until it reaches an upper limit $\tau_{max}$. By using a monotonically increasing $\tau_k$ for the penalty term in (3.15a), the algorithm reduces the slack variables $s_i$ and $\hat{s}_i$ very quickly. As a result, new iterates quickly become feasible as $s_i$ and $\hat{s}_i$ vanish. The upper limit $\tau_{max}$ is imposed to avoid numerical difficulties that may occur if $\tau_k$ becomes too large and to ensure convergence

if a feasible region is not found [9]. Consequently, while formulation (3.15) accepts *infeasible* initial points, the iterates obtained by solving (16) are practically identical to those obtained by solving (3.14).

### 3.2.4 The Algorithm

The input parameters for the algorithm include the bounds $\delta_i$ on the measurement error. Setting $\delta_i$ to a lower value leads to a "tighter" solution. On the other hand, a larger $\delta_i$ would make the algorithm less sensitive to outliers. However, some a priori knowledge about noise statistic, if available, or sensor geometry can be used to derive reasonable values for $\delta_i$. For example, if measurement noise $\boldsymbol{\varepsilon}$ obeys a Gaussian distribution with zero mean and known covariance $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2)$, then $\delta_i$ can be expressed as $\delta_i = \gamma\sigma_i$, where $\gamma$ is a parameter that determines the width of the confidence interval. For example, for $\gamma = 3$ we have the probability $Pr\{|\varepsilon_i| \leq 3\sigma_i\} \approx 0.99$. Other input parameters are initial point $\boldsymbol{x}_0$, maximum number of iterations $K_{max}$, initial weight $\tau_0$, and upper limit of weight $\tau_{max}$ (to avoid numerical problems that may occur if $\tau_i$ becomes too large).

As mentioned in Sec. 3.2 of the thesis, the original LS objective is highly nonconvex with many local minimums even for small-scale systems. Consequently, it is of critical importance to select a good initial point for the proposed PCCP-based algorithm because PCCP is essentially a local procedure. Several techniques are available, these include:

(i) Select the initial point uniformly randomly over the same region as the unknown radiating source;

(ii) Set the initial point to the origin;

(iii) Run the algorithm from a set of candidate initial points and identify the solution as the one with lowest LS error. Typically, comparing the results from $n$ distinct initial points shall suffice. For the planar case ($n = 2$), for example, it is sufficient to compare the two intersection points of the two circles that are associated with the two smallest distance readings as the target is very likely to be in the vicinity of these sensors;

(iv) Apply a global localization algorithm such as those in [15] to generate an approximate LS solution, then take it as the initial point to run the proposed

algorithm.

The algorithm can be now outlined as follows.

**Algorithm 3. PCCP-based LS Algorithm for Source Localization**

1) Input data: Sensor locations $\{\boldsymbol{a}_i, i = 1, \ldots, m\}$, range measurements $\{r_i, i = 1, \ldots, m\}$, initial point $\boldsymbol{x}_0$, maximum number of iterations $K_{max}$, initial weight $\tau_0$ and upper limit of weight $\tau_{max}$, weight increment $\mu > 0$, error bounds $\delta_i$. Set iteration count to $k = 0$.

2) Form $\boldsymbol{v}_k$ as

$$\boldsymbol{v}_k = \bar{\boldsymbol{a}} + \frac{1}{m}\sum_{i=1}^{m} r_i \partial \|\boldsymbol{x}_k - \boldsymbol{a}_i\|, \quad \bar{\boldsymbol{a}} = \frac{1}{m}\sum_{i=1}^{m} \boldsymbol{a}_i$$

and solve

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{s},\hat{\boldsymbol{s}}}{\text{minimize}} \quad & \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{v}_k + \tau_k \sum_{i=1}^{m}(s_i + \hat{s}_i) \\ \text{subject to:} \quad & \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i - \delta_i \le s_i \\ & -\|\boldsymbol{x}_k - \boldsymbol{a}_i\| - \frac{(\boldsymbol{x}_k - \boldsymbol{a}_i)^T}{\|\boldsymbol{x}_k - \boldsymbol{a}_i\|}(\boldsymbol{x} - \boldsymbol{x}_k) + r_i - \delta_i \le \hat{s}_i \\ & s_i \ge 0, \hat{s}_i \ge 0, \text{ for: } i = 1, 2, \ldots, m \end{aligned}$$

Denote the solution as $(\boldsymbol{s}^*, \hat{\boldsymbol{s}}^*, \boldsymbol{x}^*)$.

3) Update $\tau_{k+1} = \min(\mu\tau_k, \tau_{max})$, set $k = k + 1$.

4) If $k = K_{max}$, terminate and output $\boldsymbol{x}^*$ as the solution; otherwise, set $\boldsymbol{x}_k = \boldsymbol{x}^*$ and repeat from Step 2.

## 3.3 Numerical Results

For illustration purposes, the proposed algorithm was applied to a network with five sensors, and its performance was evaluated and compared with existing state-of-the-art methods by Monte Carlo simulations with a set-up similar to that of [15]. SR-LS solutions were used as performance benchmarks for the PCCP-based LS Algorithm. The system consisted of 5 sensors $\{\boldsymbol{a}_i, i = 1, 2, \ldots, 5\}$ randomly placed in the planar region in $[-15; 15] \times [-15; 15]$, and a radiating source $\boldsymbol{x}_s$, located randomly in the region $\{\boldsymbol{x} = [x_1; x_2], -10 \leq x_1, x_2 \leq 10\}$. The coordinates of the source and sensors were generated for each dimension following a uniform distribution. Measurement noise $\{\varepsilon_i, i = 1, \ldots, m\}$ was modelled as independent and identically distributed (i.i.d) random variables with zero mean and variance $\sigma^2$, with $\sigma$ being one of four possible levels $\{10^{-3}, 10^{-2}, 10^{-1}, 1\}$. The range measurements $\{r_i, i = 1, 2, \ldots, 5\}$ were calculated using (3.1). Accuracy of source location estimation was evaluated in terms of average of the squared position error error in the form $\|\boldsymbol{x}^* - \boldsymbol{x}_s\|^2$, where $\boldsymbol{x}_s$ denotes the exact source location and $\boldsymbol{x}^*$ is its estimation obtained by SR-LS and PCCP methods, respectively. In our simulations parameter $\gamma$ was set to 3 and the number of iterations was set to 20. The proposed method was implemented by using CVX [45] and implementation of SR-LS followed [15]. The PCCP algorithm was initialized with intersection points of the two circles that are associated with the two smallest distance readings. A candidate solution point with lowest LS error in (3.2) was chosen as a PCCP solution. In cases when the circles did not intersect due to high noise level, the initial point was set as a midpoint between the centers of the two circles.

Table 3.1 provides comparisons of the PCCP with SR-LS and MLE, where each entry is averaged squared error over 1,000 Monte Carlo runs of the method. The MLE was implemented using Matlab function *lsqnonlin* [46], initialized with the same point as PCCP. It is observed that, comparing with SR-LS, the estimates produced by the proposed algorithm are found to be closer to the true source locations in MSE sense. The last column of the table represents relative improvement of the proposed method over SR-LS solutions in percentage.

Figure 3.2 illustrates the bar chart of location estimation errors for the SR-LS solution and the proposed algorithm for various numbers of sensors. Each entry is a squared error of the method averaged over 50 random initializations of the system consisting of $m$ sensors, with $m$ being one of $\{4, 5, 7, 10\}$.

Table 3.1: Averaged MSE for SR-LS and PCCP methods

| $\sigma$ | MLE | SR - LS | PCCP | R.I. |
|---|---|---|---|---|
| 1e-03 | 6.0159e-01 | 1.3394e-06 | **9.5243e-07** | 29% |
| 1e-02 | 3.5077e-01 | 1.4516e-04 | **9.5831e-05** | 34% |
| 1e-01 | 3.7866e-01 | 1.2058e-02 | **8.7107e-03** | 28% |
| 1e+0 | 1.4470e+00 | 1.3662e+00 | **1.2346e+00** | 10% |

Finally, we study the convergence of the PCCP-based LS Algorithm. As an example, consider an instance of the source localization problem on the plane ($n = 2$) as discussed in Sec. 2.1 of the thesis. The system consists of five sensors ($m = 5$) located at $(6, 4)^T, (0, -10)^T, (5, -3)^T, (1, -4)^T$ and $(3, -3)^T$ with the source emmitting the signal at $\boldsymbol{x}_s = (-2, 3)^T$. Figure 3.3 shows the iteration path the PCCP-based LS follows when initialized with a randomly selected point $\boldsymbol{x}_0 = (-20, -10)^T$. It takes 8 iterations for the PCCP-based LS to converge to a point $\boldsymbol{x}^* = (, )^T$ which is a good estimate of the R-LS approximation of the exact source location $\boldsymbol{x}_s$.

Figure 3.3 depicts the convergence speed of the proposed approach for 50 random initializations of the system with given number of sensors ($m$). In simulations, we set $\sigma = 10^{-1}$ and consider systems with 4, 5, 7 and 10 sensors randomly placed in the planar region in $[-15; 15] \times [-15; 15]$. For every estimate of the PCCP-based LS, we compute the value of the objective in (3.6). It is observed that PCCP-based LS converges fast, in most cases after 10 iterations. For systems with large numbers of sensor nodes ($m = 10$) PCCP-based LS converges approximately in three sequential updatings.

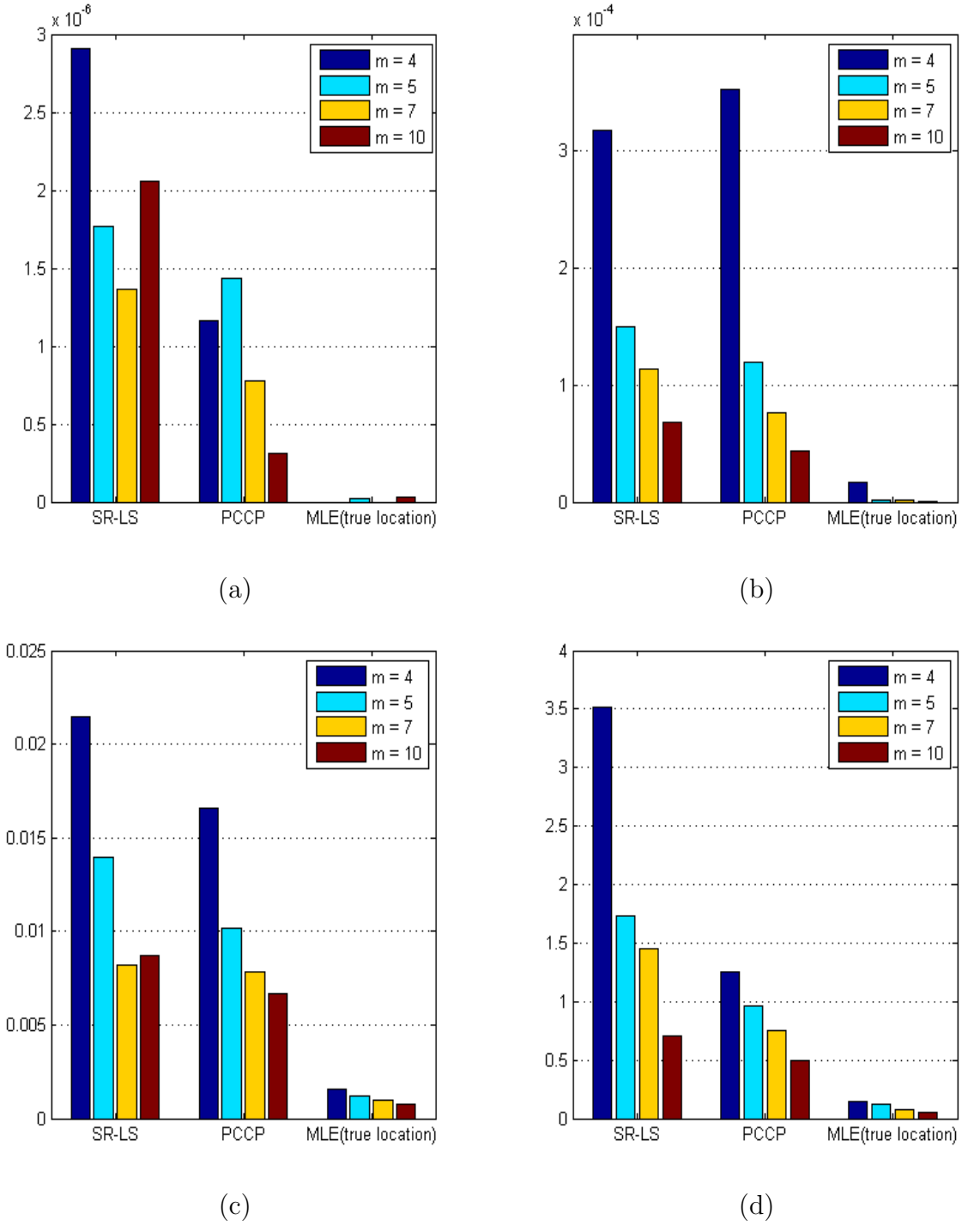



Figure 3.2: MSE for different methods and various number of sensor nodes $m$ and different noise levels with (a) $\sigma = 10^{-3}$, (b) $\sigma = 10^{-2}$, (c) $\sigma = 10^{-1}$, and (d) $\sigma = 1$.

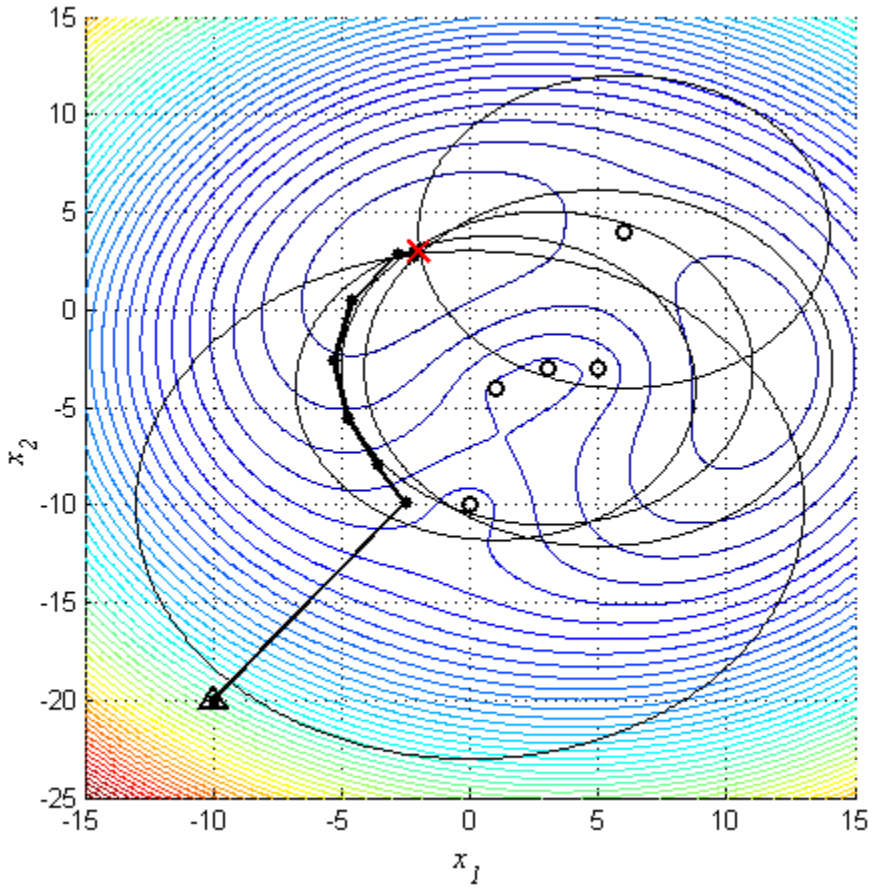



Figure 3.3: Iteration path of the PCCP-based LS Algorithm and contours of the R-LS objective function over the region $\Re = \{\boldsymbol{x} : -15 \leq x_1 \leq 15, -25 \leq x_2 \leq 15\}$. The red cross indicates the location of the signal source. Sensors are located at $(6, 4)^T, (0, -10)^T, (5, -3)^T, (1, -4)^T$ and $(3, -3)^T$. Large circles denote possible source locations given the range reading at a particular sensor.

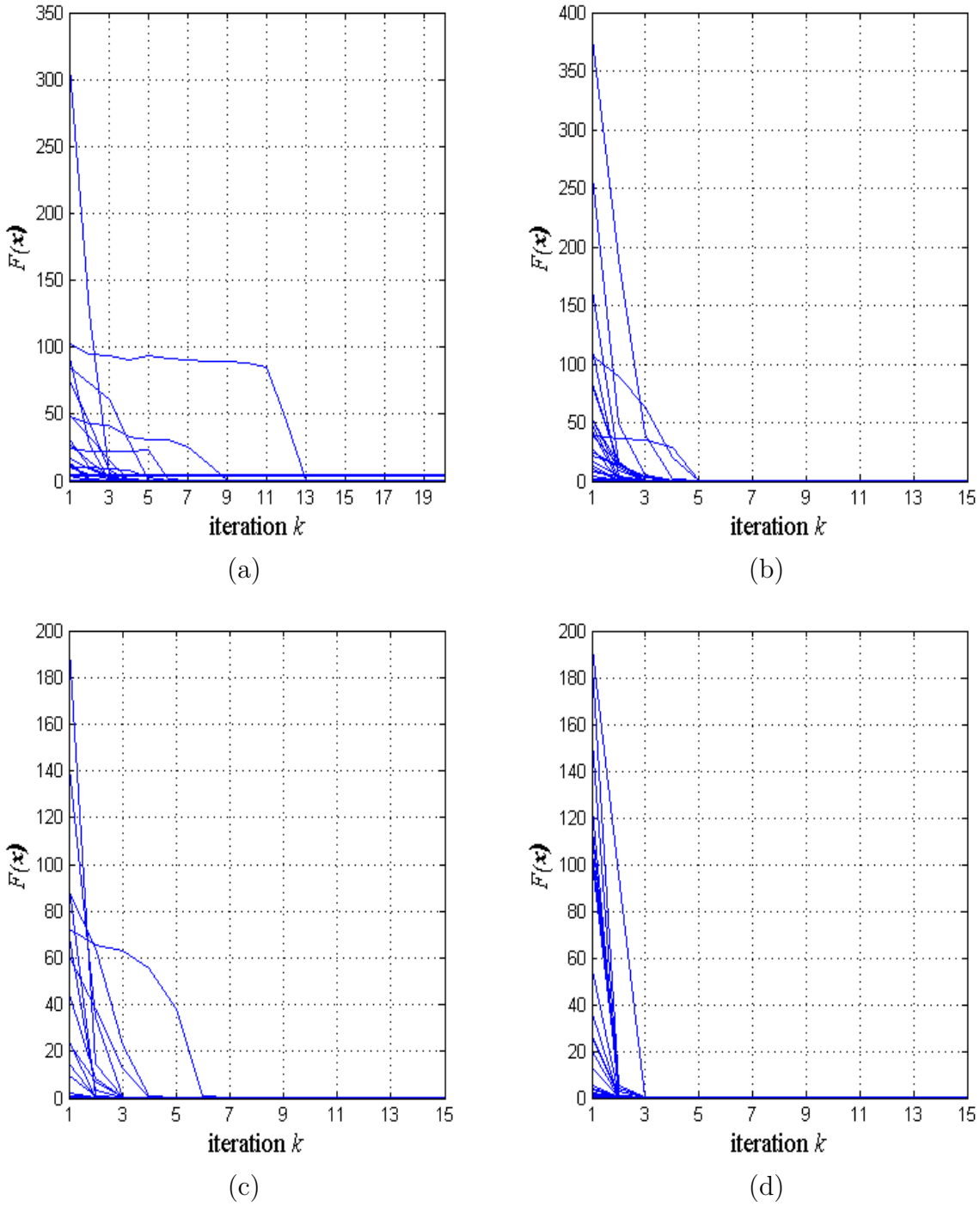



Figure 3.4: Convergernce of the proposed PCCP-based LS for random initializations with $\sigma = 10^{-1}$ for (a) 4 sensor nodes, (b) 5 sensor nodes, (c) 7 sensor nodes, and (d) 10 sensor nodes.

# Bibliography

[1] J. O. Smith and J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. on Acoustic, Speech Signal Proc.*, vol. 12, pp. 1661–1669, Dec. 1987.

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

[3] K. Yao, R. Hudson, C. Reed, D. Chen, and F. Lorenzelli, "Blind beamforming on a randomly distributed sensor array system," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 1555-1567, Oct. 1998.

[4] M. A. Sprito, "On the accuracy of cellular mobile station location estimation," *IEEE Trans. on Veh. Technol.*, vol. 50, pp. 674-685, May 2001.

[5] Y. Huang, J. Benesty, G. W. Elko, and R. M. Mersereau, "Realtime passive source localization: A practical linear correction least-squares approach," *IEEE Trans. on Speech Audio Proc.*, vol. 9, no. 8, pp. 943-956, Nov. 2002.

[6] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Trans. on Signal Proc.*, vol. 52, no. 4, pp. 1121–1228, Apr. 2004.

[7] D. Li and H. Hu, "Energy-Based Collaborative Source Localization Using Acoustic Microsensor Array," in *EURASIP Journal on Applied Signal Proc.*, vol. 4, 321–337, 2003.

[8] X. Sheng and Y.-H. Hu, "ML Multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Trans. on Signal Proc.*, vol. 53, no.1, pp. 44–53, Jan. 2005.

[9] Z. M. Saric, D. D. Kukolj, and N. D. Teslic, "Acoustic source localization in wireless sensor network", *Circuits Syst Signal Proc.*, vol. 29, pp. 837–856, 2010.

[10] L.Lu, H.-C. Wu, K.Yan, and S.Iyengar, "Robust expectation maximization algorithm for multiple wideband acoustic source localization in the presence of nonuniform noise variances," *IEEE Sensors J.*, vol. 11, no. 3, pp. 536–544, Mar. 2011.

[11] K.W. Cheung, W.K. Ma, and H.C. So, "Accurate approximation algorithm for TOA-based maximum-likelihood mobile location using semidefinite programming," in *Proc. ICASSP*, vol. 2, pp. 145–148, 2004.

[12] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location," *IEEE Signal Pro. Mag.*, vol. 22, no. 4, pp. 24–40, July 2005.

[13] Y. T. Chan, H. Y. C. Hang, and P. C. Ching, "Exact and approximate maximum likelihood localization algorithms," *IEEE Trans. on Veh. Technol.*, vol. 55, no. 1, pp. 10–16, Jan. 2006.

[14] P. Stoica and J. Li, "Source localization from range-difference measurements," *IEEE Signal Proc. Mag.*, vol. 23, pp. 63–65,69, Nov. 2006.

[15] A. Beck, P. Stoica and J. Li, "Exact and approximate solutions of source localization problems," *IEEE Trans. on Signal Proc.*, vol. 56, no. 5, pp. 1770–1777, May 2008.

[16] A. Beck, M. Teboulle, and Z. Chikishev, "Iterative minimization schemes for solving the single source localization problem," *SIAM J. Optim.*, vol. 19, no. 3, pp. 1397–1416, Nov. 2008.

[17] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Comm. Pure Appl. Math.*, vol. 63, pp. 1–38, 2010.

[18] A. Beck, and D. Pan, "On the solution of the GPS localization and circle fitting problems," *SIAM J. Optim.*, vol. 22, no. 1, pp. 108–134, Jan. 2012.

[19] A. Beck, "On the convergence of alternating minimization for convex programming with applications to iteratively re-weighted least squares and decomposition schemes," *SIAM J. Optim.*, vol. 25, no. 1, pp. 185-209, Jan. 2015.

[20] J.J. More, "Generalizations of the trust region subproblem," *Optim. Methods Softw.*, vol. 2, pp. 189–209, 1993.

[21] C. Fortin and H. Wolkowicz, "The trust region subproblem and semidefinite programming," *Optim. Methods Softw.*, vol. 19, no.1, pp. 41–67, 2004.

[22] D.P. O'Leary, "Robust regression computation using iteratively reweighted least squares," *SIAM J. Matrix Anal. Appl.*, vol. 11, no. 3, pp. 466–480, 1990.

[23] N. Bissantz, L. Dümbgen, A. Munk, and B. Stratmann, "Convergence analysis of generalized iteratively reweighted least squares algorithms on convex function spaces," *SIAM J. Optim.*, vol. 19, no. 4, pp 1828–1845, 2009.

[24] K. W. Cheung and H. C. So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE Trans. on Signal Proc.*, vol. 53, no. 2, pp. 460–470, Feb. 2005.

[25] S. Qin, Q. Wan, and L.-F. Duan, "Fast and efficient multidimentional scaling algorithm for mobile positioning," *IET Signal Processing*, vol. 6, no. 9, pp. 857–861, March 2012.

[26] J.A. Costa, N. Patwari, and A. O. Hero, "Distributed weighted-multidimentional scaling for node localization in sensor networks,"ACD Trans. Sens. Netw., vol. 2, no. 1, pp. 39–64, 2006.

[27] H.C. So and F.K.W. Chan, "Generalized Subspace Approach for Mobile Positioning With Time-of-Arrival Measurements," *IEEE Trans. on Signal Proc.*, vol. 55, no. 10, pp. 5103–5107, October 2007.

[28] H. Liu, H. Darabi, P. B, and J. Liu, "Survey of Wireless Indoor Positioning Techniques and Systems," *IEEE Trans. on Systems, Man and Cybernetics. Part C: Applications and Revies*, vol. 37, no. 6, pp. 1067–1080, Nov. 2007.

[29] Y. Liu, Y.H. Hu, and Q. Pan, "Distributed, Robust Acoustic Source Localization in a Wireless Sensor Network," *IEEE Trans. on Signal Proc.*, vol. 60, no. 8, pp. 4350–4359, Aug. 2012.

[30] W. Kim, J. Lee, and G. Jee, "The interior-point method for an optimal treatment of bias in trilateration location," *IEEE Trans. Veh. Technol.*, vol. 55, no. 4, pp. 1291–1301, Jul. 2006.

[31] T. Qiao, S. Redfield, A. Abbasi, Z. Su, and H. Liu, "Robust coarse position estimation for TDOA localization," *IEEE Wireless Commun. Lett.*, vol. 2, no. 6, pp. 623–626, Dec. 2013.

[32] A. L. Yuille and A. Rangarajan, "The concave-convex procedure,' *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.

[33] T. Lipp and S. Boyd, "Variations and extensions of the convex-concave procedure," *Research Report*, Stanford University, Aug. 2014.

[34] G. R. Lanckreit and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Advances in Neural Information Proc. Systems*, pp. 1759–1767, 2009.

[35] Y. Kang, Z. Zhang, and W.-J. Li, "On the global convergence of majorization minimization algorithms for nonconvex optimization problems," in *arXiv: 1504.07791v2*, 2015.

[36] M. R. Gholami, S. Gezici, and E. G. Ström, "TW-TOA based positioning in the presence of clock imperfections," *Digital Signal Processing*, vol. 59, pp. 19–30, 2016.

[37] M. R. Gholami, E. G. Ström, F. Sottile, D. Dardari, A. Conti, S. Gezici, M. Rydström, and M. A. Spirito, "Static positioning using UWB range measurements," in *Future Network and Mobile Summit Conference Proceedings*, pp. 1–10, 2010.

[38] D. Dardari, A.Conti, U.J. Ferner, A. Giorgetti, and M.Z. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proceedings of the IEEE*, vol. 97, pp. 427–450, 2009.

[39] M. R. Gholami, E. G. Ström, and M. Rydström, "Indoor sensor node positioning using UWB range measurements," in *17th European Signal Processing Conference (Eusipco)*, pp. 1943–1947, 2009.

[40] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 40–95, Mar. 1996.

[41] P. Hartman, "On functions representable as a difference of convex functions," *Pacific Journal of Math*, vol. 9, no. 3, pp. 707–713, 1959.

[42] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, Springer, New-York, 2007.

[43] Y. Nesterov, *Introductory Lectures on Convex Optimization*, Kluwer Academic Publishers, Boston, 2004.

[44] C. Gentile, N. Alsindi, R. Raulefs, and C. Teolis, *Geolocation Techniques: Principles and Applications*, Springer, New-York, 2013.

[45] CVX Research, http://cvx.com/cvx, August 2012.

[46] The Mathworks Inc., http://mathworks.com, 2015.

[47] "Iterative re-weighting"

[48] "Penalty Convex-Concave Procedure of Source Localization Problem,"