Localization Algorithms in Passive Sensor Networks

by

Darya Ismailova
B.Sc., 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 in Passive Sensor Networks

by

Darya Ismailova
B.Sc., 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 Tables

List of Figures

ACKNOWLEDGEMENTS

Chapter 1

Introduction

Tips from Uvic template:

1.1 My Claims

State what's new here. Make an impact here. How about something like the following box:

I make *four* claims which my dissertation validates:

> My new algorithm to solve the problem of doing nothing include these important new features whose practical applicability can be proved both formally and empirically:
>
> 1. first feature;
> 2. second feature;
> 3. everything is much easier to understand, and therefore, easier to implement correctly.

Claim 1 and claim 2 are *quantitative* - they will be proven by experiment.
Claim 3 is *qualitative* - they will be demonstrated by argument.

1.1.1 The Importance of My Claims

Some very important positive consequences arise from the validation of the above claims. It is these consequences that comprise a significant positive contribution to research in the field of whatever the field is.

Claim 1 implies that:

1. Something profound which applies to:
    - something excellent;
    - something important.
2. Something else just as profound.

Claim 2 implies that:

- Repeat as above if necessary and useful.

The consequence of claim 3 is that:

- There must be something good coming out of all this work!

1.2 Agenda

Chapter 1 contains a statement of the claims which will be proved by this dissertation followed by an overview of the structure of the document itself.

Chapter 2 describes in details the open problem which is to be tackled together with its context, its impact and the overall motivation for the research overall.

Chapter 3 gives the new research, its methodology, the algorithms involved, the new solution, the new work done. Formal proofs and arguments are made here. This is the first of the two contributions expected in a thesis for a graduate degree.

Chapter 4 is where the experiments and the methodology for them is fully described. The first part includes all details of how the empirical side of the research has been conducted. Note that not every thesis has this empirical portion.

Chapter 5 includes the evaluation of the data presented above and the comparisons with the work of others, to show how much better the new approach is. This is the second of the two contributions expected in a thesis for a graduate degree. Note that this part could be consolidated into the chapter above.

Chapter 6 contains a restatement of the claims and results of the dissertation. It also enumerates avenues of future work for further development of the concept and its applications.

1.3 The Problem to be Solved

TODO

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

TOA,

TDOA,

AOA ?,

non-range-based?

Chapter 2

Improved Least-Squares Methods for Source Localization: An Iterative Re-Weighting Approach

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] - [14]. 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 [27], however most of them are *local* methods that are iterative, hence extremely 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 [10], where the least squares model is relaxed to a semidefinite programming problem which is known to be convex [26], hence robust to where it starts. Another representative in this class is reference [14], 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 [14] 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 the 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, on comparing with the best known results, is found to be considerably closer to the original range-based (range-difference-based) LS solution.

2.1 Source Localization From 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 ith 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 ε_i denotes the unknown noise that has occurred when the ith 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] -[14]. 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 ε_i are independent identically distributed (Gaussian) variables with zero mean and variance σ_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 there are many methods for continuous unconstrained optimization [27], 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

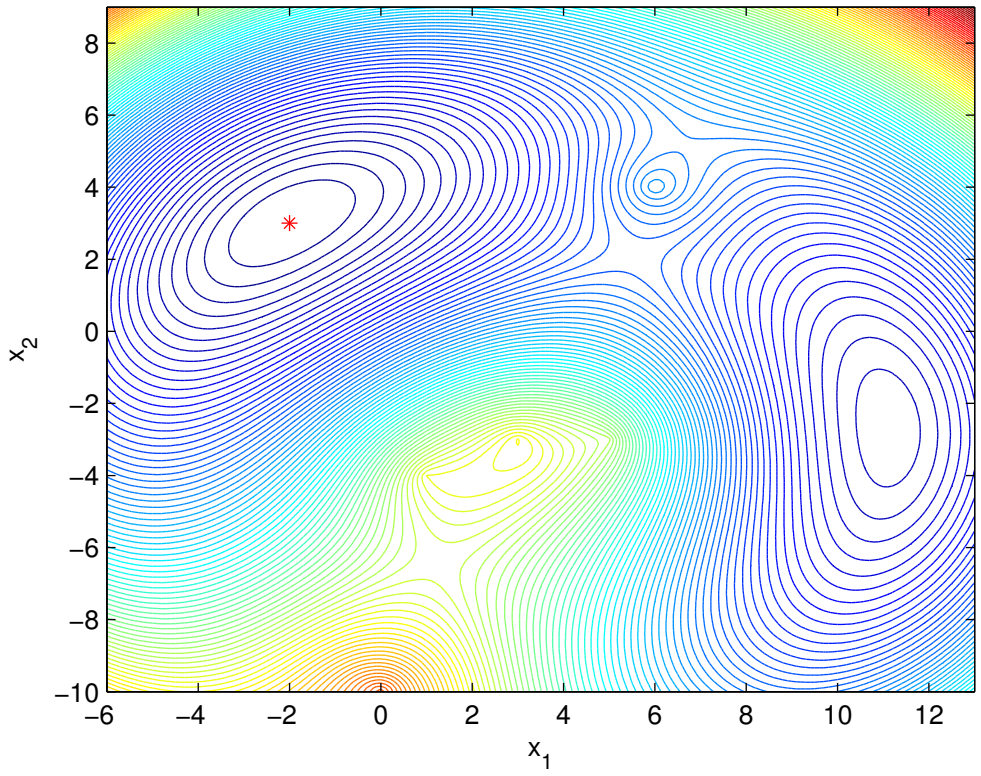

Figure 2.1: Contours of the R-LS objective function over the region $\Re = \{\boldsymbol{x} : -6 \leq x_1 \leq 13, -10 \leq x_2 \leq 9\}$

method gives no guaranty about its 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) describes 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 [10] addresses problem (2.2) by a convex relaxation technique where

(2.2) is modified to a convex problem known as semidefinite programming (SDP) [26]. 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 [26, 27]. 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 [14] 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 α 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 $\boldsymbol{y} = [\boldsymbol{x}^T \|\boldsymbol{x}\|^2]^T$ and

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

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

This problem conversion, made in [14], 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) [18, 19] whose global solutions can be computed by exploring the KKT conditions which are both necessary and sufficient optimality conditions in the case of GTRS [18].

Few remarks are now in order. 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 [14] 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 [15] and [16], 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 {*cite*}. The method has found many applications, such as in robust regression {*cite*}, sparse recovery {*cite*}, 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 {*cite*}. This problem can be stated as: given m points $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m \in R^n$ called *ancors* 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 m ancors:

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

Fermat–Weber problem is much easier to analyze and solve than the ML problem (2.2) since it is a well-structured nonsmooth convex minimization problem, however the similarities of these problems has been noted in the literature [15]. The standard fixed point (SFP) algorithm proposed in [15] is a gradient method with a fixed step size. However, being a gradient method, it has chances to converge to local minimum. The sequential weighted least squares algorithm (SWLS) proposed in [15] is also an iterative method where each iteration involves solving a nonlinear least squares problem similar to (2.9). As discussed in [15] the SWLS method is superior over SFP in terms of convergence rate and a wider region of convergence to the global minimum. However, it still does have a chance to converge to a local minimum in case of certain sensor setup even if an initial point $\boldsymbol{x}_0$ is constructed by following a procedure developed specifically for this purpose. The method presented below is different from the above described approaches in a 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 a weighted SR-LS (WSR-LS) problem, namely,

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

Following [14], 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 $\boldsymbol{\Gamma} = \text{diag}\left(\sqrt{w_1}, \ldots, \sqrt{w_m}\right)$. 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 ith 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

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

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 kth 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 kth 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 ζ.

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

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

via (2.16)

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

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

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. 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 (2.19) 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. This section presents 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.

The well-known Newton algorithm [27] 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 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 δ a small positive constant. Obviously, $\tilde{\boldsymbol{H}}(\boldsymbol{x})$ is guaranteed to be positive definite. The search direction in the kth iteration of the modified Newton algorithm is given by

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

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

σ	SR - LS	IRWSR-LS (Im.,%)	*hybrid* IRWSR-LS (Im.,%)
1e-03	2.0325e-06	1.1996e-06 (41)	1.1994e-06 (41)
1e-02	1.8372e-04	1.2480e-04 (32)	1.2481e-04 (32)
1e-01	1.8361e-02	1.2223e-02 (33)	1.2214e-02 (33)
1e+0	2.3752e+00	1.5108e+00 (36)	1.5330e+00 (35)

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 [14]. SR-LS and SRD-LS solutions were used as performance benchmarks for Algorithm 1 and its variant. In both cases the system consisted of m sensors $\{\boldsymbol{a}_i, i = 1, 2, \ldots, m\}$ randomly placed in the planar region in R^2, and a radiating source $\boldsymbol{x}_s$, located randomly in the region R=$\{\boldsymbol{x} : (x_1, x_2)^T, -10 \leq x_1, x_2 \leq 10\}$. 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 σ^2. 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. 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: MSE(% Improvement).

Employing a set-up similar to that in [14], the simulation studies of Algorithm 1 and its variant considered $m = 5$ sensors placed in the region $[-15; 15] \times [-15; 15]$, with σ 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 (2.1) and Step 4 of Algorithm 1 was imple-

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

σ	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
1e+0	1.1920e+00	4.3266e+00	4.3266e+00

mented using the SR-LS algorithm proposed in [14]. 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 suggest that, again, IRWSR-LS and hybrid IRWSR-LS outperform SR-LS. Figures (2.2 - 2.5) describe 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 σ being one of $\{10^{-3}, 10^{-2}, 10^{-1}, 1\}$. Here $\boldsymbol{x}^*$ denotes the estimated location and $\boldsymbol{x}_s$ is the exact location of the source. Histograms that correspond to the results obtained by IRWSR-LS are shifted closer towards 0 than those obtained by SR-LS and have smaller variance. Also, the solutions obtained by running IRWSR-LS have fewer outliers with large error.

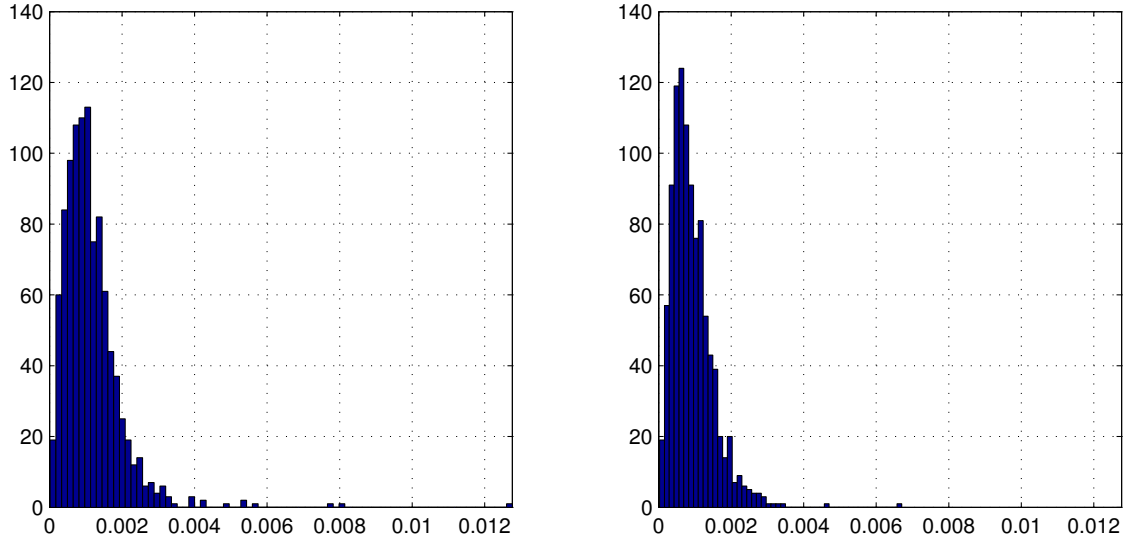

Figure 2.2: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, noise $\sigma = 10^{-3}$

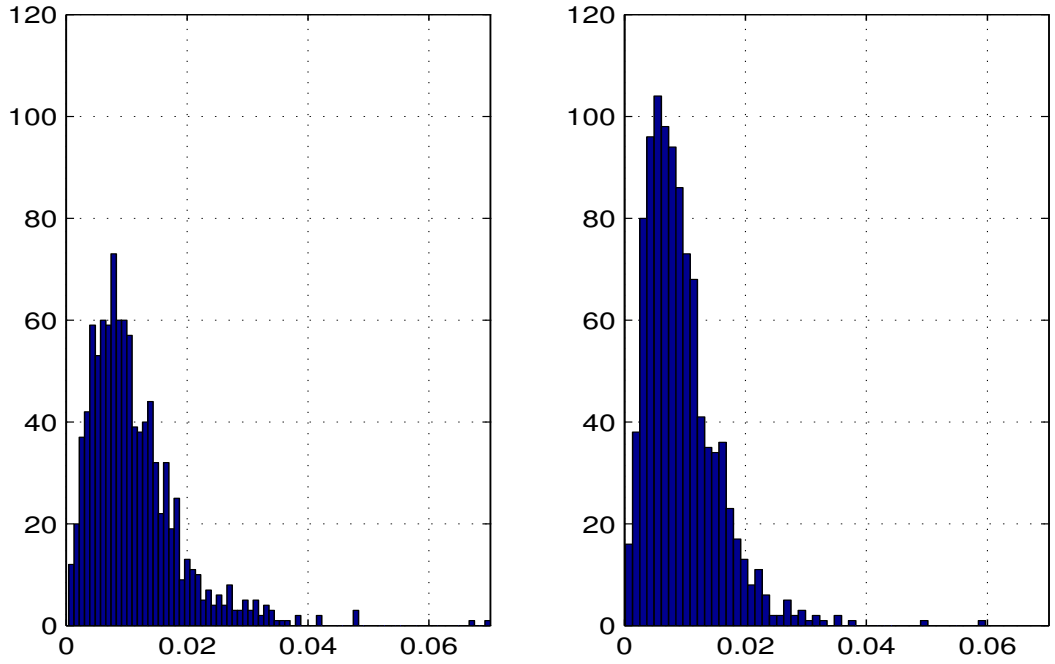

Figure 2.3: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, noise $\sigma = 10^{-2}$

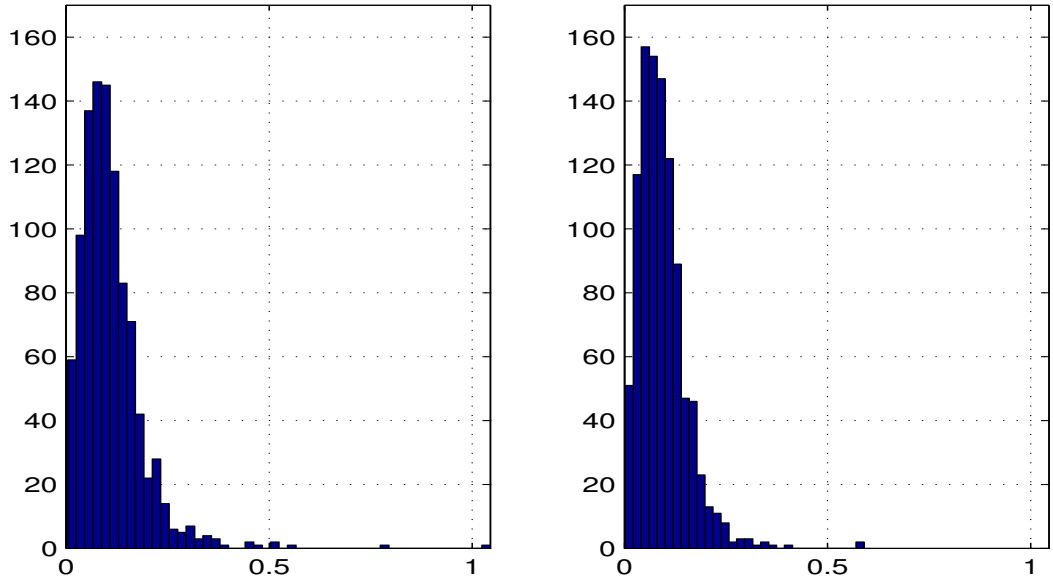

Figure 2.4: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, noise $\sigma = 10^{-1}$

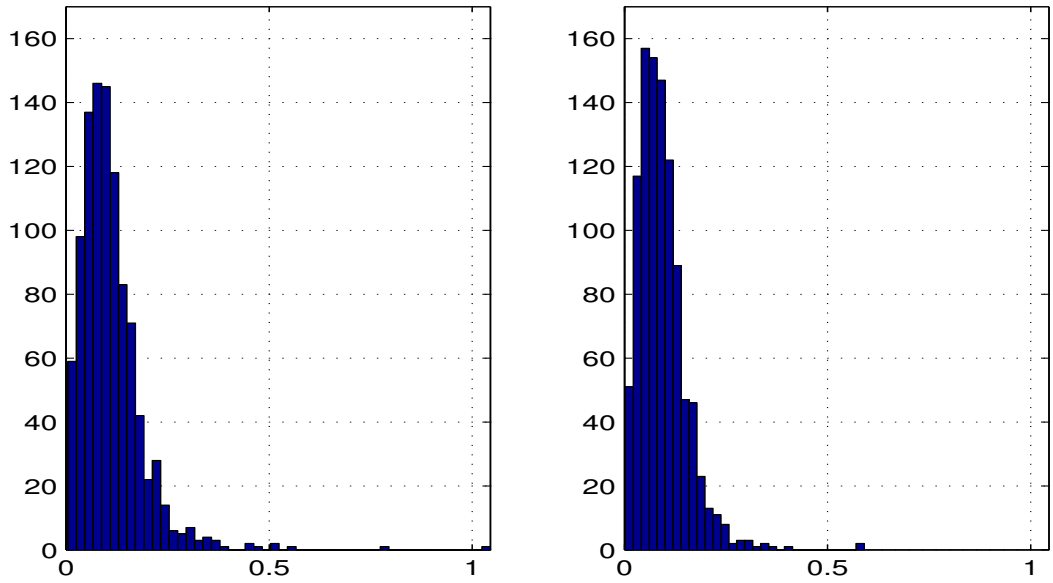

Figure 2.5: Histograms of the errors of the SR-LS (left) and IRWSR-LS (right) solutions, noise $\sigma = 1$

2.2 Source Localization From 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 [13, 14].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) technique is based on the idea that the position of the mobile device 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 syncronization between mobile devices is not available. Each TDOA measurement constrains the location of the mobile device to be on a hyperboloid with a constant range-difference between the two reference points. A TDOA measurement between base station (sensor) i and j can be given by*

$$t_{ij} = (t_i - t_x) - (t_j - t_x) = t_i - t_j$$

where t_x is the clock time of the mobile device, t_i and t_j are the TOA between the device and stations i and j respectively. The equation can be written in terms of distance through speed of light scaling or

$$d_{ij} = (t_i - t_j)c = d_i - d_j = \|\boldsymbol{a}_i - \boldsymbol{x}\| - \|\boldsymbol{a}_j - \boldsymbol{x}\|$$

Without loss of generality, the latter equation is valid with the assumption that the station j is placed at the origin of the coordinate system. [21] The idea of TDOA is to determine the relative position of the mobile transmitter by examining the difference in time at shich the signal arrives at multiple measuring units, rather than the absolute arrival time. As usual, these sensors are denoted as $\{\boldsymbol{a}_0, \boldsymbol{a}_1, \ldots, \boldsymbol{a}_m\} \in R^n$ with $\boldsymbol{a}_0 = \boldsymbol{0}$ be placed at the origin and used as a *reference sensor*. The range-difference r_i is defined as the difference between the distance from sensor $\boldsymbol{a}_i$ to source $\boldsymbol{x}$ and the distance from sensor $\boldsymbol{a}_0$ to source $\boldsymbol{x}$, namely

$$r_i = \|\boldsymbol{x} - \boldsymbol{a}_i\| - \|\boldsymbol{x} - \boldsymbol{a}_0\| = \|\boldsymbol{x} - \boldsymbol{a}_i\| - \|\boldsymbol{x}\|, i = 1, \ldots, m \tag{2.21}$$

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

$$d_i = r_i + \varepsilon_i, \text{ for } i = i = 1, 2, \ldots, m \tag{2.22}$$

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

$$\underset{\boldsymbol{x}\in R^n}{\text{minimize}}\, F(\boldsymbol{x}) = \sum_{i=1}^{m} \left(d_i + \|\boldsymbol{x}\| - \|\boldsymbol{x} - \boldsymbol{a}_i\|\right)^2 \tag{2.23}$$

Unfortunately, finding the global solution of (2.23) turns out to be a very hard problem. **include example of nonconvexity : a contour and a surface plots**. Reference [14] proposes a squared range-difference LS (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

$$\left(d_i + \|\boldsymbol{x}\|\right)^2 = \|\boldsymbol{x} - \boldsymbol{a}_i\|^2 \tag{2.24}$$

which can be simplified to

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

where $g_i = d_i^2 - \|\boldsymbol{a}_i\|^2$. In practice (2.25) does not hold exactly due to measurement noise that contaminates the data d_i's. In other words, if d_i's in (2.25) are taken to be real-world data, then we only have

$$-2d_i\|\boldsymbol{x}\| - 2\boldsymbol{a}_i^T\boldsymbol{x} - g_i \approx 0, i = 1, \ldots, m \tag{2.26}$$

Reference [14] proposes a LS solution for the problem at hand by minimizing the sum of squared residues on the left side of (2.26), namely,

$$\underset{\boldsymbol{x}\in R^n}{\text{minimize}} \sum_{i=1}^{m} \left(-2\boldsymbol{a}_i^T\boldsymbol{x} - 2d_i\|\boldsymbol{x}\| - g_i\right)^2 \tag{2.27}$$

By introducing new variable $\boldsymbol{y} = [\boldsymbol{x}^T \|\boldsymbol{x}\|]^T$ and noticing nonnegativity of the compo-

nent y_{n+1} problem (2.27) is converted to

$$\underset{\boldsymbol{y}\in R^{n+1}}{\text{minimize}} \|\boldsymbol{B}\boldsymbol{y}-\boldsymbol{g}\|^2 \tag{2.28a}$$

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

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

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

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

Because of the presence of the nonnegativity constraint in (2.28c), (2.28) is no longer a GTRS problem hence the technique used for the case of range measurements does not apply. Nevertheless reference [14] presents a rigorous argument which shows that the optimal solution of (2.28) either assumes the form of

$$\tilde{\boldsymbol{y}}(\lambda)=\left(\boldsymbol{B}^T\boldsymbol{B}+\lambda\boldsymbol{C}\right)^{-1}\boldsymbol{B}^T\boldsymbol{g} \tag{2.28}$$

where λ 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 (2.28a), 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 (2.28) 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 SRD-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 (2.28) 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 $\boldsymbol{B}_w = \boldsymbol{\Gamma}\boldsymbol{B}$, $\boldsymbol{g}_w = \boldsymbol{\Gamma}\boldsymbol{g}$ and $\boldsymbol{\Gamma} = \text{diag}\{\sqrt{w_1}, \ldots, \sqrt{w_m}\}$, which will be referred to as the weighted SRD-LS (WSRD-LS) problem. On comparing (2.31) with (2.28), it follows immediately that the global solver for problem (2.28) characterized by data set $\{\boldsymbol{B}, \boldsymbol{g}, \boldsymbol{C}\}$ can also be sued 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 (2.24), (2.25) and observe that the ith 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 ith term of the objective function in the RD-LS problem (2.23) 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 kth 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 ξ.

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 on how to solve (2.31) can be found in Appendix 1.

A variant of Algorithm 2

Like in 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 (2.23), 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 its 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 (2.23).

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,

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

may be 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 λ_{min} be the smallest eigenvalue of $\boldsymbol{H}(\boldsymbol{x})$. If λ_{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

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

Table 2.3: Averaged MSE for SRD-LS and IRWSRD-LS methods

σ	SRD - LS	IRWSRD-LS	Improvement, %
1e-04	8.4918e-09	4.1050e-09	51.7
1e-03	5.8553e-06	3.5105e-06	40.0
1e-02	6.3508e-05	5.0378e-05	20.7
1e-01	1.6057e-02	1.0055e-02	37.3
1e+0	1.2773e+00	6.2221e-01	51.2

where $\tilde{\mathbf{\Lambda}} = \text{diag}\left(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_n\right)$ 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 δ 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 [14] by Monte Carlo simulations with a set-up similar to that of [14]. SRD-LS solutions were used as performance benchmarks for Algorithm 2 and its variant. In both cases the system consisted of m sensors $\{\boldsymbol{a}_i, i = 1, 2, \ldots, m\}$ randomly placed in the planar region in R^2, 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 σ^2. Accuracy of source location estimation was evaluated in terms of mean squared error in the form 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 10,000 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: MSE(% Improvement).

The simulation studies of Algorithm 2 considered $m = 11$ sensors $\{\boldsymbol{a}_i, i = 1, 2, \ldots, 10\}$ with $\boldsymbol{a}_0 = \boldsymbol{0}$ and other ten sensors placed in the region $[-15; 15] \times [-15; 15]$, with σ being one of five possible levels $\{10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1\}$. The range-difference measurements used to form matrix $\boldsymbol{B}$ in Step 2 of the Algorithm 2 were calculated as a noise-contaminated version of d_i in (2.21) as:

$$d_i^{noisy} = d_i + \varepsilon_i.$$

Step 4 of Algorithm 2 was carried out using the SRD-LS algorithm in [14]. Again, the IRWSRD-LS solutions offer considerable improvement over SRD-LS solutions.

2.3 Extensions

As noted in [13], methods developed in this chapter for localization based on range measurements can be adopted to solve the problem of sourse localization from energy measurements.

The theory supporting this statement can be found in [7, 8, 9, 10, 20].

The energy, acoustic or RF, of the signal received by the sensor is inversely proportional to the distance between sensor and the radiating source ([7, 8, 9, 10, 20]).

"The localization based on the received signal strength uses the property that sound energy attenuates with the square of the distance from the source."

This section provides an example of ...

In this section, a reformulation of the maximum liklihood source localization using acoustic energy meausrements is offered.

(from [13]) "The source localization problem from range measurements is related to the problem of source localization from energy measurements [7, 8, 9, 10, 20]. The energy measurement based source localization approach, advocated in [7], [9] is based on the fact that the energy of the signal received by the ith sensor over a (relatively small) time interval is inversely proportional to $\|\boldsymbol{x} - \boldsymbol{a}_i\|$, for $i = 1, 2, \ldots, m$.

Using this fact and some simple manipulations (see, e.g., [7] for details), it is possible to obtain an equation in the unknown vector x that is somewhat similar to (2), namely:....."

Acoustic energy attenuation model presented here is based on assumptions of [7] and [9] (or please refer for details). Only single sourse localization is investigated in this section.

2.3.1 Acoustic Energy Attenuation Model and Parameter Estimator

Let m be a number of acoustic sensors. For consistency of notation, let $\boldsymbol{a}_i$ denote the known location of the sensor i in space R^n, $n = 2$ 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. (**OR** Received signal is an acoustic pulse M-samples wide.) 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 ith sensor and the source. g_i is a factor that takes into account ith 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. S is the unknown acoustic energy measured 1 unit distance away from the source. α is the energy decay factor and is usually assumed to have a value 2 [7], [29].ε_i denotes the square of the background noise affecting the measurement of sensor i. Based on the central limit theorem, it can be approximated well as a normal distribution/ It is assumed to be a wide-sense stationary Gaussian random process [22] / , namely, $\varepsilon_i \sim N(\mu_i, \sigma_i^2)$ with a positive mean value $\mu_i > \sigma_i$ that is no less that the standard deviation σ_i that can be estimated empirically from data samples. For justification and validity of this energy attenuation model, please see [8], [7], [9], [22] and its references.

References [8], [7] argues that the maximum likelihood estimation of the vector of unknown parameters $\boldsymbol{\theta} = [\boldsymbol{x}^T S]^T$ can be obtained my minimizing the quadratic form

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

where

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

and $\boldsymbol{Z}$ are (estimated) normalized energy measurements for the case of the single radiating source. Important assumption about 2.34 is that μ_i and σ_i are considered known.

2.3.2 Reformulation

There are two things to note about (2.34). First, (2.34) is 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, although there are m sensors reporting the acousting energy reading in fact there are $n+1 \leq m$ unknowns, including the unknown acoustic energy S radiated from the source. To eliminate the unknown source energy S [8] propose first to compute the ratio k_{ij} of the calibrated energy readings from ith and jth 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 a 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 ρ_{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 $\boldsymbol{a}_j$:

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

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 $\forall\,\{i,j\} \in I_1 : 1 \leq i \leq m-1, i+1 \leq j \leq m$ it is true that $0 < k_{ij} \neq 1$ and $\forall\,\{i,j\} \in I_2 : 1 \leq i \leq m-1, i+1 \leq j \leq m\ k_{ij} = 1$ 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 criterion that is equivalent to (2.34):

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

For the brevity of notation the double indexes ij were replaced by single indexes l_1 and l_2. The unconstrained problem in (2.38), due to mathematical analogy with (2.9), can be solved using Algorithm 1 developed in section 3.1. After some simple manupilations and necessary variable changes, can be converted to the constrained problem that is similar to (2.10):

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

$$\text{subject to:} \quad \boldsymbol{y}^T \boldsymbol{D} \boldsymbol{y} + 2\boldsymbol{f}^T \boldsymbol{y} = 0 \tag{2.40b}$$

where

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{x} \\ \|\boldsymbol{x}\|^2 \end{bmatrix}, \boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix}, \boldsymbol{b} = \begin{pmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \end{pmatrix} \tag{2.41}$$

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

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

$$\boldsymbol{A}_1 = \begin{pmatrix} -2\boldsymbol{c}_1^T & 1 \\ \vdots & \vdots \\ -2\boldsymbol{a}_{L_1}^T & 1 \end{pmatrix}, \boldsymbol{b}_1 = \begin{pmatrix} \rho_1^2 - \|\boldsymbol{c}_1\|^2 \\ \vdots \\ \rho_{L_1}^2 - \|\boldsymbol{c}_{L_1}\|^2 \end{pmatrix} \tag{2.42}$$

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

Chapter 3

Penalty Convex-Concave Procedure for Source Localization Problem

Locating a radiating source from range 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. Least squares (LS) based algorithms for source localization problems constitute an important class of solution techniques as they are geometrically meaningful and often provide low complexity solution procedures with competitive estimation accuracy [12]-[14]. On the other hand, the error measure in an LS formulation for the localization problem of interest is shown to be highly non-convex, possessing multiple local solutions with degraded performance. This non-convexity excludes many local methods that are iterative, hence extremely sensitive to where the iteration begins. Several non-iterative *global* localization techniques are available from the literature. A global solution may be obtained by relaxing the LS model at hand to a semidefinite programming (SDP) problem which is known to be convex [26]. In doing so, however, the convexification based solution is no longer optimal in LS sense. Another representative in this class is the method proposed in [14], where localization problems for range measurements are addressed by developing solution methods for *squared* range LS (SR-LS) problems. Although these methods are efficient in terms of complexity, they remain to be suboptimal in the maximum likelihood (ML) sense because the solutions produced are merely approximations of the ML estimate.

In this paper, we focus on LS formulation for the problem of localizing a single radiating source based on range measurements. We exploit special structure of the cost function of an unconstrained LS formulation and show that it is well suited for being investigated in a setting known as difference-of-convex-functions (DC) programming. Further, we present an algorithm for solving the LS problem at hand based on a penalty convex-concave procedure (PCCP) [24] that accommodates infeasible initial points. We also provide algorithmic details that are tailored to the localization problem at hand, these include additional constraints that enforce the algorithms iteration path towards the LS solution and strategies to secure good initial points for the algorithm. 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

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 n coordinates of the ith 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

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

where ε_i denotes the unknown noise that has occurred when the ith 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 $\boldsymbol{r}$. For the localization problem at hand, the range-based least squares (R-LS) estimate refers to the solution of the problem

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

Formulation (3.2) is connected to the maximum-likelihood (ML) location estimation that determines $\boldsymbol{x}$ by examining the probabilistic model of the error vector $\boldsymbol{\varepsilon}$. If $\boldsymbol{\varepsilon}$ obeys a Gaussian distribution with zero mean and covariance $\boldsymbol{\Sigma} = \text{diag}(\sigma_1^2, \ldots, \sigma_m^2)$,

then the maximum likelihood (ML) location estimator in this case is known to be

$$\boldsymbol{x}_{ML} = \arg\min_{\boldsymbol{x}\in R^n} (\boldsymbol{r}-\boldsymbol{g})^T \Sigma^{-1} (\boldsymbol{r}-\boldsymbol{g}) \tag{3.3}$$

where $\boldsymbol{g} = [g_1\ g_2 \ldots\ g_m]^T$ with $g_i = \|\boldsymbol{x} - \boldsymbol{a}_i\|$. It follows immediately that the ML solution in (3.3) is identical to the R-LS solution of problem (3.2) when covariance $\boldsymbol{\Sigma}$ is proportional to the identity matrix, i.e., $\sigma_1^2 = \ldots = \sigma_m^2 = 1$. In the literature this is known as the equal noise power case. For notation simplicity this paper 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 (3.2) with $\{\sigma_i^{-2}, i = 1, \ldots, m\}$ as the weights.

There are many methods for continuous unconstrained optimization [27], however most of them are *local* methods in the sense they are sensitive to the choice of initial point, and give no guarantee to yield global solutions when applied to non-convex objective functions. Unfortunately, the objective function in (3.2) is highly non-convex, possessing many local minimizers even for small-scale systems. In this paper we present an different approach to solve the positioning problem, which employs a successive convex-concave procedure.

3.2 Fitting the Localization Problem to the CCP Framework

3.2.1 Basic Convex-Concave Procedure

The CCP refers to an effective heuristic method to deal with a class of *nonconvex* problems of the form

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}) - g(\boldsymbol{x}) && \text{(3.4a)} \\ \text{subject to:} \quad & f_i(\boldsymbol{x}) \le g_i(\boldsymbol{x}) \quad \text{for: } i = 1, 2, \ldots, m && \text{(3.4b)} \end{aligned}$$

where $f(\boldsymbol{x}), g(\boldsymbol{x}), f_i(\boldsymbol{x}), g_i(\boldsymbol{x})$ for $i = 1, 2, \ldots, m$ are convex. 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

$$\hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) = g(\boldsymbol{x}_k) + \bigtriangledown g(\boldsymbol{x}_k)^T(\boldsymbol{x} - \boldsymbol{x}_k) \tag{3.5a}$$

and

$$\begin{aligned} \hat{g}_i(\boldsymbol{x}, \boldsymbol{x}_k) = g_i(\boldsymbol{x}_k) + \bigtriangledown g_i(\boldsymbol{x}_k)^T(\boldsymbol{x} - \boldsymbol{x}_k) \\ \text{for: } i = 1, 2, \ldots, m \end{aligned} \tag{3.5b}$$

(ii) Solving the convex problem

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & f(\boldsymbol{x}) - \hat{g}(\boldsymbol{x}, \boldsymbol{x}_k) && \text{(3.6a)} \\ \text{subject to:} \quad & f_i(\boldsymbol{x}) - \hat{g}_i(\boldsymbol{x}, \boldsymbol{x}_k) \leq 0 && \text{(3.6b)} \\ & \text{for: } i = 1, 2, \ldots, m \end{aligned}$$

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 (4) [24]. The basic CCP requires a *feasible* initial point $\boldsymbol{x}_0$ (in the sense that $\boldsymbol{x}_0$ satisfies (6b) 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 [13].

3.2.2 Problem Reformulation

We begin by re-writing the objective function in (2) up to a constant as:

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

The objective in (7) 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

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

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

$$\begin{aligned} f(\boldsymbol{x}) &= m\boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\sum_{i=1}^{m}\boldsymbol{a}_i \\ g(\boldsymbol{x}) &= 2\sum_{i=1}^{m} r_i\|\boldsymbol{x}-\boldsymbol{a}_i\| \end{aligned} \tag{3.8}$$

the objective in (7) 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 (4a). Note that $g(\boldsymbol{x})$ in (8) 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 (5a) by a subgradient [28] 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 (5a) 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

$$c = -2\sum_{i=1}^{m} r_i \boldsymbol{a}_i^T \partial\|\boldsymbol{x}_k - \boldsymbol{a}_i\|.$$

It follows that up to a multiplicative factor $1/m$ and an additive constant term the convex objective function in (6a) can be written as

$$\underset{\boldsymbol{x}}{\text{minimize}} \quad \hat{F}(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{v}_k \tag{3.9}$$

where

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

It is rather straightforward to see that given $\boldsymbol{x}_k$ (in the k-th iteration) the solution of the quadratic problem (9) can be obtained as

$$\boldsymbol{x}_{k+1} = \bar{\boldsymbol{a}} + \frac{1}{m}\sum_{i=1}^{m} r_i \partial \|\boldsymbol{x}_k - \boldsymbol{a}_i\| \tag{3.11}$$

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

$$\left(\|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i\right)^2 \leq \delta_i^2 \tag{3.12}$$

which leads to

$$\|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i - \delta_i \leq 0 \tag{3.13a}$$

$$r_i - \delta_i \leq \|\boldsymbol{x} - \boldsymbol{a}_i\| \tag{3.13b}$$

for $1 \leq i \leq m$. The constraints in (13a) are convex and fit into those in (6b) with $f_i(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}_i\| - r_i - \delta_i$ and $g_i(\boldsymbol{x}) = 0$, while those in (13b) are in the form of (4b) with $f_i(\boldsymbol{x}) = r_i - \delta_i$ and $g_i(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}_i\|$. Following CCP (see (5b)), $g_i(\boldsymbol{x}) = \|\boldsymbol{x} - \boldsymbol{a}_i\|$ 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 (13b) 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 (6b), or equivalently

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

We remark that constraint (14) is not only convex but also tighter than (13b). 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 (14) automatically satisfies (13b). Summarizing, the convexified problem in the k-th iteration can be stated as

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{minimize}} \quad & \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{v}_k && \text{(3.15a)} \\ \text{subject to:} \quad & \|\boldsymbol{x}-\boldsymbol{a}_i\| - r_i - \delta_i \leq 0 && \text{(3.15b)} \end{aligned}$$

$$-\|\boldsymbol{x}_k-\boldsymbol{a}_i\| - \partial\|\boldsymbol{x}_k-\boldsymbol{a}_i\|^T(\boldsymbol{x}-\boldsymbol{x}_k) + r_i - \delta_i \leq 0 \qquad \text{(3.15c)}$$

A technical problem making the formulation in (15) 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 (15b) and (15c) 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:

$$\begin{aligned} \underset{\boldsymbol{x},\boldsymbol{s},\hat{\boldsymbol{s}}}{\text{minimize}} \quad & \boldsymbol{x}^T\boldsymbol{x} - 2\boldsymbol{x}^T\boldsymbol{v}_k + \tau_k\sum_{i=1}^{m}(s_i+\hat{s}_i) && \text{(3.16a)} \\ \text{subject to:} \quad & \|\boldsymbol{x}-\boldsymbol{a}_i\| - r_i - \delta_i \leq s_i && \text{(3.16b)} \end{aligned}$$

$$\begin{aligned} -\|\boldsymbol{x}_k-\boldsymbol{a}_i\| - \frac{(\boldsymbol{x}_k-\boldsymbol{a}_i)^T}{\|\boldsymbol{x}_k-\boldsymbol{a}_i\|}(\boldsymbol{x}-\boldsymbol{x}_k) + r_i - \delta_i \leq \hat{s}_i && \text{(3.16c)} \\ s_i \geq 0, \hat{s}_i \geq 0, \text{ for: } i = 1,2,\ldots,m && \text{(3.16d)} \end{aligned}$$

where the weight $\tau_k \geq 0$ increases as iterations proceed until it reaches an upper limit τ_{max}. By using a monotonically increasing τ_k for the penalty term in (16a), 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 τ_{max} is imposed to avoid numerical difficulties that may occur if τ_k becomes too large and to ensure convergence if a feasible region is not found [9]. Consequently, while formulation (16) accepts *infeasible* initial points, the iterates obtained by solving (16) are practically identical to those obtained by solving (15).

3.2.4 The Algorithm

The input parameters for the algorithm include the bound δ_i on the measurement error. Setting δ_i to a lower value leads to a "tighter" solution. On the other hand, a larger δ_i would make the algorithm less sensitive to outliers. 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 δ_i can be expressed as $\delta_i = \gamma\sigma_i$, where γ is a parameter that determines the width of 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 τ_0, and upper limit of weight τ_{max} (to avoid numerical problems that may occur if τ_i becomes too large).

As mentioned in Sec. 2, the original LS objective is highly non-convex 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; and (iv) Apply a global localization algorithm such as those in [14] 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.

PCCP-based LS Algorithm for Source Localization

Step 1: Input sensor locations $\{\boldsymbol{a}_i, i = 1, \ldots, m\}$, range measurements $\{r_i, i = 1, \ldots, m\}$, $\boldsymbol{x}_0, K_{max}, \tau_0, \tau_{max}, \mu > 0, \gamma, \sigma$, and set $k = 0$.

Step 2: Form $\boldsymbol{v}_k$ as in (10) and solve (16). Denote the solution as $(\boldsymbol{s}^*, \hat{\boldsymbol{s}}^*, \boldsymbol{x}^*)$.

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

Step 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.

Appendix A

Appendix 1

A.1 Solving 2.21

3.21

A.2 Solving 2.32c

3.32

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. D. Li and H. Hu, "Least square solutions of energy based acoustic source localization problems," in *Proc. ICPPW*, 2004.

[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] 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.

[11] 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.

[12] 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.

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

[14] 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.

[15] 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.

[16] 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.

[17] 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.

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

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

[20] 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.

[21] 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

[22] 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.

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

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

[25] 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.

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

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

[28] Y. Nesterov, *Introductory Lectures on Convex Optimization*, Kluwer, 2004.

[29] *Geolocation Techniques*

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

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