Improved Time of Arrival measurement model for non-convex optimization

Juri Sidorenko [ab*], Volker Schatz [a], Leo Doktorski [a],
Norbert Scherer-Negenborn [a], Michael Arens [a],
Urs Hugentobler [b]

[a]Fraunhofer Institute of Optronics, System Technologies and Image Exploitation IOSB, Germany - juri.sidorenko@iosb.fraunhofer.de
[b]Institute of Astronomical and Physical Geodesy, Technical University of Munich, Germany - urs.hugentobler@tum.de

Keywords: time of arrival, non-linear optimization, localisation, navigation

1 Abstract

The quadratic system provided by the Time of Arrival technique can be solved analytically or by non-linear least squares minimization. An important problem in quadratic optimization is the possible convergence to a local minimum, instead of the global minimum. This problem does not occur for global navigation satellite system (GNSS), due to the known satellite positions. In applications with unknown positions of the reference stations, such as indoor localization with self-calibration, local minima are an important issue. This article presents an approach showing how this risk can be significantly reduced. The main idea of our approach is to transform the local minimum to a saddle point by increasing the number of dimensions. In addition to numerical tests we analytically prove the theorem and the criteria that no other local minima exists for non-trivial constellations.

2 Introduction

In position estimation the Time of Arrival (ToA) [1] technique is standard. The area of applications extends from satellite-based systems like GPS [2], GLONASS [3], Galileo [4], mobile phone localization (GSM) [5], radar-based systems such as UWB [6], FMCW radar [7] to acoustic systems [8].
The ToA technique leads to a quadratic equation. Optimization algorithms used to solve this system depend on the initial estimate. Unfortunately, chosen initial estimates can cause to that the optimization algorithm converges to

the local minimum. With known reference stations positions, it is possible to transform the quadratic to a linear system [9, 10, 11]. This linear system can be used to provide an initial estimate. On the other hand, the linear system is more affected by noise, compared to the quadratic system [9, 10]. In general local minima are not an issue in applications with known reference stations positions such as GNSS. This changes if it is necessary to obtain the locations of the reference stations without additional measuring equipment, which is also known as self-calibration. The most robust self-calibration solution with noise is non-linear optimization [12]. This solution suffers if the initial estimates are not close to the global minimum [13, 14, 15]. In [16, 17] it was proposed to use semi-definite relaxation (SDP) as an initialization for the maximum likelihood (ML) estimator. Nuclear norm-based methods such as [18, 19, 20] also reduce the risk of being trapped in a local minimum. Alternatively, non-iterative methods can be used. A two-dimensional non-iterative method was proposed for the case with three transponders and three receivers in [21]. The solution for the three-dimensional case was the subject of the investigation in [22, 23]. The authors provided a non-iterative solution for the cases with (5, 5), (6,4) (10,4) transponders and receivers. The roles of the base stations and the transponders are equivalent, hence it does not matter if six base stations and four transponders are used or vice versa. In the case that one of the base station positions coincides with the position of one of the transponders, it is possible to obtain a closed form solution [24, 25]. An alternative approach is called far field [26]. If the distances between the base stations and the transponders are considerably larger than those between the base stations, it also possible to use a linear simplification. Real measurement data is highly non-convex and non-linear optimization still provides the most robust solutions [12].
We present a new approach which does not require an initial estimate at all. The idea is that an additional dimension in the l^2 norm transforms the local minimum of the ToA equation to a saddle point without adding more local minimum. This is not equivalent to a receiver time offset, which would be added or subtracted outside the norm. Under the constraint that the position of the reference stations is known, we prove that with our approach the local minima becomes a saddle point and no further local minima exist for non-trivial constellations. This paper focuses on proving our approach. Further publications will be based on this concept and will investigate its practical use for self-calibration. Our approach is based on introducing an additional dimension. It can be seen as a kind of lifting method, a generic term covering many numerical methods introducing an additional variable. Lifting methods have been used for solving non-linear optimization problems [27], machine learning [28]problems, optimal control problems [29], boundary value problems [30] and parameter estimation problems in ordinary differential equations (ODE) [31]. To the best of our knowledge, our specific approach has not been applied before, in particular in the context of ToA localization. It does not work with every objective function but we will show that it is suitable for ToA localization equations.

This paper is organized as follows. The next section, introduces the objective functions F and the corresponding improved objective functions F_L. In Section

4, we use Levenberg-Marquardt algorithm [32] to illustrate the optimization steps for F and F_L. The last section address the results of the optimization algorithm with randomly selected constellations.

3 Methodology

Table 1: Used notations

Notations	Definition
x, y, z	Estimated position of object T
x_G, y_G, z_G	Ground truth position of object T
a_i, b_i, c_i	Ground truth position of base stations B_i, $1 \leq i \leq N$
d_i	Distance measurements between base stations B_i and object T
λ	Additional variable

Figure 1 shows three base stations B_i at known positions (a_i, b_i, c_i), and one object T at unknown position (x, y, z). The distances measurements d_i between base stations B_i and object T are known and derived from time of flight measurements. The unknown position of object T can be estimated by the known positions of the base stations B_i and the distance measurements d_i. Measurement errors are neglected in this paper, therefore distances measurements can be referred to as distances.

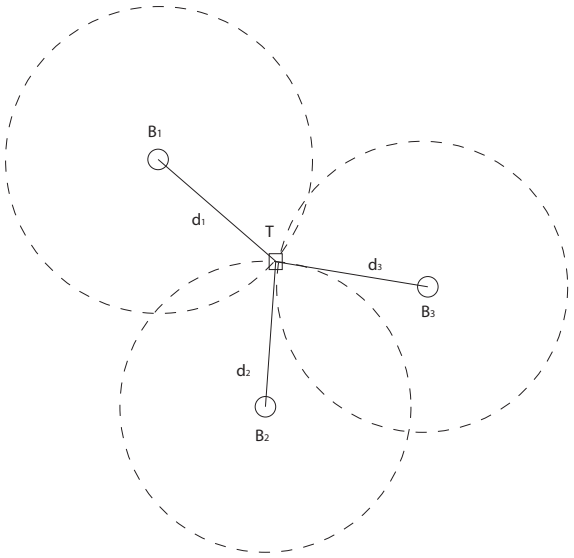

Figure 1: The dashed circles are the distances between the base stations B_i and object T. The object T is located at the intersection point between the three dashed circles.

3.1 Mathematical formulation

The distances between the base station B_i and object T are defined as

$$d_i^2 = (x_G - a_i)^2 + (y_G - b_i)^2 + (z_G - c_i)^2. \qquad 1 \leq i \leq N$$

- Objective function F_1:

$$F_1(x, y, z) \coloneqq \frac{1}{4} \sum_{i=1}^{N} \left[\sqrt{(x - a_i)^2 + (y - b_i)^2 + (z - c_i)^2} - d_i \right]^2 \tag{1}$$

- Objective function F_2:

$$F_2(x, y, z) \coloneqq \frac{1}{4} \sum_{i=1}^{N} \left[(x - a_i)^2 + (y - b_i)^2 + (z - c_i)^2 - d_i^2 \right]^2 \tag{2}$$

[2, 33]. Finding the minimum value of the cost function of (1) or (2) can be done by non-convex optimization [34] $F_i(x, y, z) \rightarrow argmin$. Alternatively, the non-linear system can be transformed into a linear system [9, 10]. With the assumptions made in this section 3.1 it is possible to obtain a linear system. In more complex cases where the positions of base stations B_i are unknown this is not possible at all. With regard to future extensions to determining the base station positions as well as the location of the object T, this article focuses on finding a solution with a non-convex optimization algorithm.

3.2 Reason for the approach

The objective functions (1) and (2) are non-linear and non-convex. The optimization of the objective functions can cause convergence to a local minimum L instead of the global minimum G. In our approach, instead of F_1 (1) and F_2 (2) the improved objective functions F_{L1} and F_{L2} are used. Both have an additional variable λ compared to the F functions.

- Improved objective function F_{L1}:

$$F_{L1}(x, y, z, \lambda) \coloneqq \frac{1}{4} \sum_{i=1}^{N} \left[\sqrt{(x - a_i)^2 + (y - b_i)^2 + (z - c_i)^2 + \lambda^2} - d_i \right]^2 \quad (3)$$

- Improved objective function F_{L2}:

$$F_{L2}(x, y, z, \lambda) \coloneqq \frac{1}{4} \sum_{i=1}^{N} \left[(x - a_i)^2 + (y - b_i)^2 + (z - c_i)^2 + \lambda^2 - d_i^2 \right]^2 \quad (4)$$

In the next section, we prove that the F_{L2} (4) has a saddle point at every position of the local minimum $L(x_L, y_L, z_L)$ of F_2 (2) . Therefore, the Levenberg-Marquardt algorithm has a lower probability of converging to a local minimum. The additional variable λ is not equivalent to a receiver time offset in GNSS, which would be added or subtracted outside the norm.

3.3 Characteristics of a local minimum

First assumption

The objective function has a unique global minimum at $G(x_G, y_G, z_G)$ and at least one local minimum at $L(x_L, y_L, z_L)$.

Second assumption

It is known that the first derivative of F_L with respect to x, y and z is zero at the local minimum. The second assumption is that the second derivative of F_L at the same position is positive (Table 2).

Table 2: Assumption two

First derivative	Second derivative
$\left(\frac{\partial}{\partial x} F_L\right)(x_L, y_L, z_L, 0) = 0$	$\left(\frac{\partial^2}{\partial x^2} F_L\right)(x_L, y_L, z_L, 0) > 0$
$\left(\frac{\partial}{\partial y} F_L\right)(x_L, y_L, z_L, 0) = 0$	$\left(\frac{\partial^2}{\partial y^2} F_L\right)(x_L, y_L, z_L, 0) > 0$
$\left(\frac{\partial}{\partial z} F_L\right)(x_L, y_L, z_L, 0) = 0$	$\left(\frac{\partial^2}{\partial z^2} F_L\right)(x_L, y_L, z_L, 0) > 0$

3.4 Assertion

The first derivative of F_L with respect to the additional variable λ is zero and the second derivative is less than zero at the local minimum (See Table 3). In combination with the first and second assumption the local minimum becomes a saddle point. The Levenberg-Marquardt (derivative based optimization algorithm) would not converge to a saddle point.

Table 3: Assertion

First derivative	Second derivative
$\left(\frac{\partial}{\partial\lambda}F_L\right)(x_L, y_L, z_L, 0) = 0$	$\left(\frac{\partial^2}{\partial\lambda^2}F_L\right)(x_L, y_L, z_L, 0) < 0$
	$\left(\frac{\partial^2}{\partial x\partial\lambda}F_L\right)(x_L, y_L, z_L, 0) = 0$
	$\left(\frac{\partial^2}{\partial y\partial\lambda}F_L\right)(x_L, y_L, z_L, 0) = 0$
	$\left(\frac{\partial^2}{\partial z\partial\lambda}F_L\right)(x_L, y_L, z_L, 0) = 0$

Assertion for function F_1

Every local minimum of function F_1 (1) becomes a saddle point at the same coordinates with function F_{L1} (3) . We have no analytical proof of this assertion, but the numerical results in Section 4 demonstrate its validity in practice.

Assertion for function F_2

Every local minimum of function F_2 (2) becomes a saddle point at the same coordinates with function F_{L2} (4) . This assertion is proven analytical in the appendix 7.1 and demonstrated numerically in 4.

3.5 The effect of an additional variable on the global minimum

At the global minimum the additional variable λ must be zero and the second derivative must be positive. The second derivative of $F_{L2}(4)$ with respect to λ at the global minimum is:

$$\left(\frac{\partial^2}{\partial\lambda^2}F_{L2}\right)(x_G, y_G, z_G, \lambda_G) = 3\lambda_G^2 N = 0\,.$$

If the second derivative is zero, a higher order derivative is required.

$$\left(\frac{\partial^3}{\partial\lambda^3}F_{L2}\right)(x_G, y_G, z_G, \lambda_G) = \sum_{i=1}^{N} 6\lambda_G N = 0$$

The third derivative is also zero. Finally, the fourth derivative is greater than zero, hence the additional variable has no effect on the global minimum.

$$\left(\frac{\partial^4}{\partial\lambda^4} F_{L2}\right)(x_G, y_G, z_G, \lambda_G) = 6N$$

3.6 No new local minima for F_{L2} with $\lambda \neq 0$

We have shown that the modified objective function F_{L2} (4) turns the local minima of F_2 (2) into saddle points and leaves the global minimum unaffected. It must still be proven that F_{L2} does not introduce new local minima that might adversely affect convergence to the global minimum.

In this section we will show that in practically relevant base station arrangements, F_{L2} has no stationary points for $\lambda \neq 0$ and $\mathbf{x} \neq \mathbf{x}_G$, and therefore no minima that would lead an optimization method astray. We will show that if the first derivative of F_{L2} with respect to λ vanishes when $\lambda \neq 0$, its gradient in the spatial directions is non-zero for $\mathbf{x} \neq \mathbf{x}_G$. This proof is best presented in vectorial notation. We will use $\mathbf{x} = (x, y, z)^T$ for the position argument and $\mathbf{a}_i = (a_i, b_i, c_i)^T$ for the base station locations.

$$\begin{aligned}
\frac{\partial}{\partial\lambda} F_{L2}(\mathbf{x}, \lambda) &= \lambda \sum_i \left((\mathbf{x} - \mathbf{a}_i)^2 + \lambda^2 - d_i^2\right) = 0, \qquad \lambda \neq 0 \\
&\Rightarrow \quad \sum_i \left((\mathbf{x} - \mathbf{a}_i)^2 + \lambda^2 - d_i^2\right) = 0 \qquad (5) \\
\operatorname{grad}_{\mathbf{x}} F_{L2}(\mathbf{x}, \lambda) &= \sum_i \left((\mathbf{x} - \mathbf{a}_i)^2 + \lambda^2 - d_i^2\right)(\mathbf{x} - \mathbf{a}_i) \qquad (6)
\end{aligned}$$

Equation (5) allows us to add or subtract any term not dependent on the summation index i in the right-hand factor of (6). We subtract $\mathbf{x}$ and add $\mathbf{a}_* = \frac{1}{N}\sum_{i=1}^N \mathbf{a}_i$, the geometrical center of the base stations:

$$\begin{aligned}
\operatorname{grad}_{\mathbf{x}} F_{L2}(\mathbf{x}, \lambda) &= \sum_i \left((\mathbf{x} - \mathbf{a}_i)^2 + \lambda^2 - d_i^2\right)(\mathbf{x} - \mathbf{a}_i - \mathbf{x} + \mathbf{a}_*) \\
&= -\sum_i \left((\mathbf{x} - \mathbf{a}_i)^2 + \lambda^2 - d_i^2\right)(\mathbf{a}_i - \mathbf{a}_*)\,.
\end{aligned}$$

By the construction of $\mathbf{a}_*$, we have $\sum_{i=1}^N (\mathbf{a}_i - \mathbf{a}_*) = 0$, so now we can add or subtract any term not depending on the summation index in the left-hand factor. We add $-\lambda^2 - \mathbf{x}^2 + \mathbf{x}_G^2$ and substitute $d_i = |\mathbf{x}_G - \mathbf{a}_i|$, expand the squares and simplify, obtaining:

$$
\begin{aligned}
\operatorname{grad}_{\mathbf{x}} F_{L2}(\mathbf{x},\lambda) &= -\sum_i \left((\mathbf{x}-\mathbf{a}_i)^2 - \mathbf{x}^2 + \mathbf{x}_G^2 - d_i^2\right)\,(\mathbf{a}_i - \mathbf{a}_*) \\
&= -\sum_i \left((\mathbf{x}-\mathbf{a}_i)^2 - \mathbf{x}^2 + \mathbf{x}_G^2 - (\mathbf{x}_G-\mathbf{a}_i)^2\right)\,(\mathbf{a}_i - \mathbf{a}_*) \\
&= \sum_i (2\,\mathbf{x}\,\mathbf{a}_i - 2\,\mathbf{x}_G\,\mathbf{a}_i)\,(\mathbf{a}_i - \mathbf{a}_*) \\
&= 2\,(\mathbf{x}-\mathbf{x}_G)^T \sum_i \mathbf{a}_i \otimes (\mathbf{a}_i - \mathbf{a}_*) \\
&= 2\,(\mathbf{x}-\mathbf{x}_G)^T\,\mathbf{M}\,.
\end{aligned}
$$

Here, $\mathbf{u} \otimes \mathbf{v}$ denotes the outer product, resulting in a matrix with the entries $u_i\,v_j$. The matrix $\mathbf{M}$ can be expressed in the following form:

$$
\begin{aligned}
\mathbf{M} &= \sum_i \mathbf{a}_i \otimes \mathbf{a}_i - \left(\sum_i \mathbf{a}_i\right) \otimes \mathbf{a}_* = \sum_i \mathbf{a}_i \otimes \mathbf{a}_i - N\,\mathbf{a}_* \otimes \mathbf{a}_* \\
&= \sum_i (\mathbf{a}_i - \mathbf{a}_*) \otimes (\mathbf{a}_i - \mathbf{a}_*)\,.
\end{aligned}
$$

The last step is analogous to the well-known derivation of the variance of a data set. The result represents $\mathbf{M}$ as a sum of unnormalized projection matrices onto the directions to the base stations from their center.

The calculation above shows that the gradient of F_{L2} has the form of a vector times a sum of projection matrices at all local minima with $\lambda \neq 0$. Projection matrices are positive semidefinite by construction, and their null space is the sub-space orthogonal to the projection direction. When adding several positive semidefinite matrices, the null space of the result is the intersection of the null spaces of the individual matrices, in our case the sub-space orthogonal to all projection directions. For unambiguous location in n (2 or 3) dimensions, at least $n+1$ base stations are needed, and they must be arranged in a non-degenerate way, i. e. so that the $\mathbf{a}_i - \mathbf{a}_*$ are a spanning set of the whole space. This makes the matrix $\mathbf{M}$ positive definite, and the gradient of F_{L2} cannot be zero for $\mathbf{x} \neq \mathbf{x}_G$. Therefore there are no local minima that prevent an optimization method from converging to the global minimum.

3.7 Two-dimensional example

In Section 7.1 it was proven that the F_{L2} (4) has a saddle point at the coordinates of the local minimum of F_2 (2) . In this section an example is created with known coordinates of the global $G(1,0)$ and local minimum $L(0,0)$. The aim of this example is to illustrate the converging steps of the Levenberg-Marquardt algorithm for the F_2 and F_{L2}. The positions of the local minimum and global minimum leads to the coordinates of base stations B_i (Table 9).

Figure 2 shows a top view perspective of the base stations B_i and object T positions inside a Cartesian coordinate system. The circles represent the distance measurements between the base stations B_i and the object T. In this example are the measurements not corrupted by noise, hence the ground truth position of object T is located at the intersection point between all circles. Nonlinear optimization algorithms can find the ground truth by minimizing the residues of the per-defined objective function to obtain the global minimum. Under certain constellations of the base stations it is possible that the used objective function has next to a global minimum a local minimum. A more detailed description of the requirements for this kind of constellation can be found in the appendix (7.2). The local and global minima have both the commonality, that the first derivative is zero and the second derivative is higher than zero. This attribute make the local minimum a trap for derivative based nonlinear optimization algorithms. In this example the local minimum is located at $L(0, 0)$ and the global minimum at $G(1, 0)$.

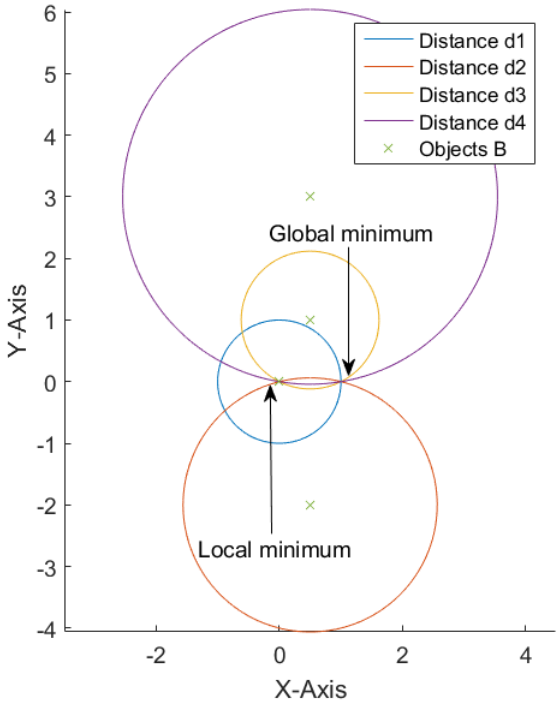

Figure 2: The circles represents the true distance between base stations B_i and the global minimum. The blue, red, yellow and magenta circles are the distances between base stations B and object T respectively.

Figure 3 shows the search space of objective function F_2 (2) and the zoom at the global minimum. The x-axis of figure 3 is the equivalent to the x-axis in figure 2 and the y-axis represent the result of objective function F_2. It can be observed that the first derivative is zero and the second derivative is higher than zero for the local minimum $L(0,0)$ and the global minimum $G(1,0)$. The only difference between both minima is the result of the objective function. At the ground truth position $G(1,0)$ the result of the objective function is zero and at the local minimum $L(0,0)$ it equates one. In case of bad initial estimates, close to the local minimum it is possible that the derivative based nonlinear optimization algorithm converges to the local minimum and remains there. The main aspect of our approach is to transform this local minima to a saddle point and eliminate this trap.

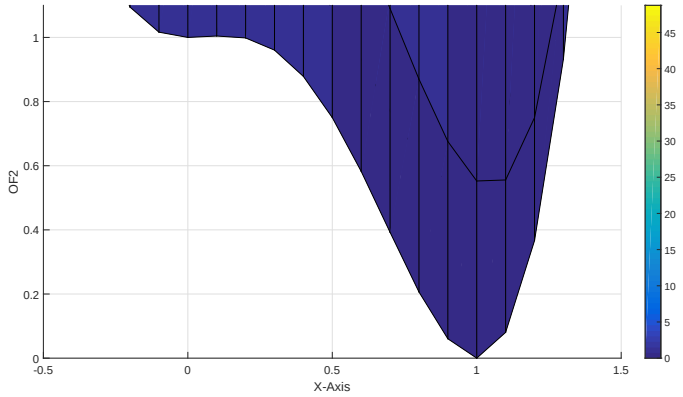

Figure 3: The Local minimum is at $L(0,0)$ and global minimum at $G(1,0)$. Colors ranging from blue to yellow show the residues of the objective function

3.7.1 Local optimization

The Levenberg-Marquardt algorithm uses the derivative to obtain the stepsize, therefore it is important that the initial estimate for the additional variable λ is non-zero. Otherwise λ remains zero, and F_{L2} (4) is effectively reduced to F_2 (2) . The initial estimates for the optimization are $x = -1$, $y = 2$ and $\lambda = 1$.

Figure 4 is equivalent to figure 2. The x-y axis represent the positions of the base stations B_i and object T inside a Cartesian coordinate system. The circles are the two dimensional euclidean distances between the base stations B_i and object T. The main difference between figure 2 and figure 4 is the additional dimension λ with the initial estimate one. The stepsize of the optimization algorithm is obtained by the first derivative of the objective function. The first derivative for the additional dimension at $\lambda = 0$ is always zero, this leads to a stepsize with the value zero.$\left(\frac{\partial}{\partial\lambda}F_2\right)(x,y,z) = \sum_{i=1}^{N}\left[(x-a_i)^2 + (y-b_i)^2 + \lambda^2 - d_i^2\right]\lambda.$

Without using the additional dimension the nonlinear optimization algorithm converges to the local minimum $L(0,0)$. On the other hand if the improved objective function F_{L2} is used, the same optimization algorithm converges to the global minimum $G(1,0)$. The optimization steps with objective function F_2 are represented by the green line and for the improved objective function F_{L2} by the blue line. It can be seen that the improved objective function F_{L2}allows the optimization algorithm to use the additional dimension of freedom to bypass the local minimum.

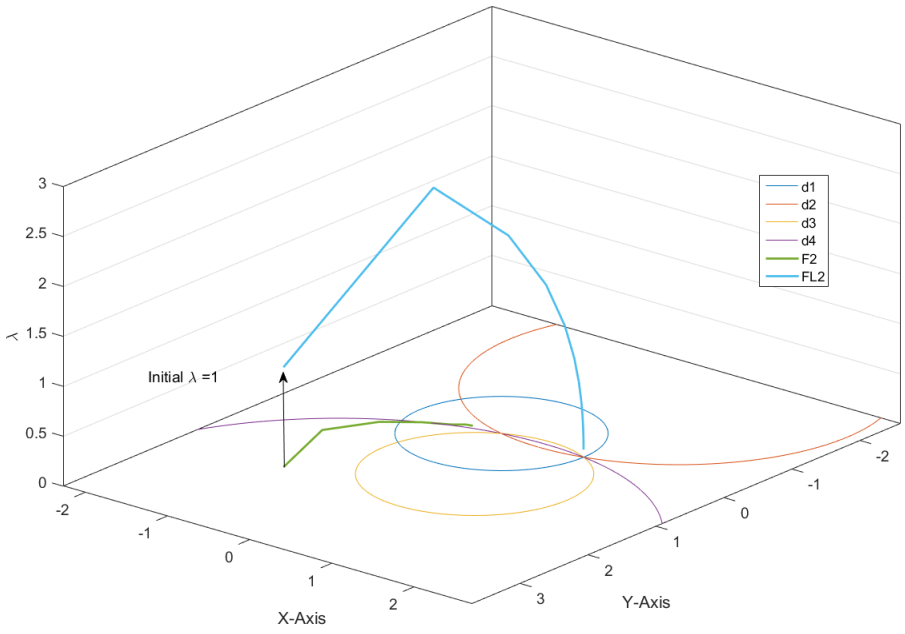

Figure 4: Iteration steps of the Levenberg-Marquardt algorithm for $F_2$and F_{L2}. F2: Objective function F_2. FL2: Improved objective function F_{L2}. Blue line: Opimization steps of F_2. Green line: Optimization steps of F_{L2}. The circles blue, red, yellow and magenta are the distances between base stations B_i and object T

Figure 5 is divided in two plots. Both plots are based on the same coordinate system as figure 2, with the same positions for base stations B_i, object T, the local $L(0,0)$ and global minimum $G(1,0)$. The colored lines represents the optimization steps. Blue lines stands for the convergence to the local minimum and the green lines to the global minimum. In the left plot it can be observed that the correct convergence for the objective function F_2 highly depends on the initial estimate of the x-coordinate. With an initial estimate $x > 0$ the

optimization algorithm converges to the global minimum, otherwise to the local minimum. In the right plot the improved objective function F_{L2} is used. At this point the initial estimates for the x axis are not relevant anymore. The optimization algorithm always converges to the global minimum $G(1, 0)$ if the initial estimate for the additional dimension $\lambda$is unequal zero.

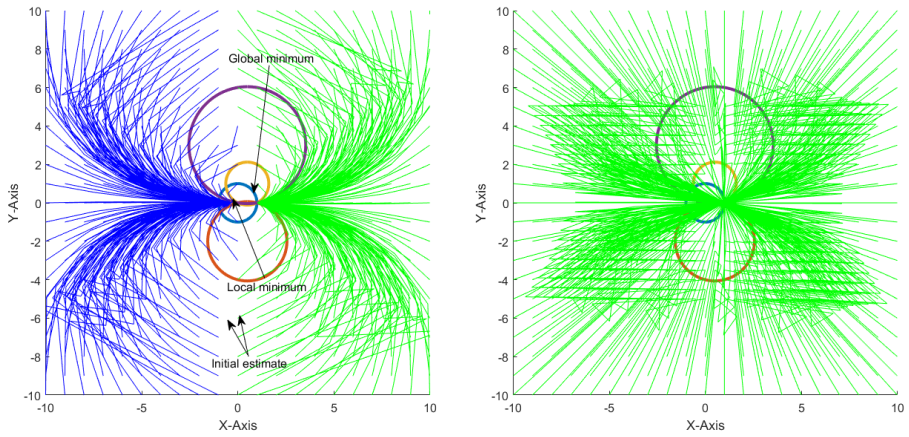

Figure 5: Left figure shows F_2 and the right figure shows F_{L2} with different initial estimates. Green: Convergence to global minimum. Blue: Convergence to local minimum.

4 Numerical results

The base stations B_i, object T and initial estimates were randomly generated by the MATLAB 'randn()' function. This function provides normally distributed random numbers in a per-defined range. This range was limited to a 10x10x10 cube. Random generated base station positions have the risk to create collinear constellations with two solutions. This constellations have been avoided by considering the normalized singular value of the co-variance matrix. This value provide information about how spread out the base stations are relative to each other. The used threshold for the normalized singular value was set to 0.1. Constellations with a higher value were rejected. The quality of the result is evaluated by the euclidean distance between the fitted value x_F, y_F and z_F and the ground truth position x_G, y_G and z_G .

- Error term:

$$E = \sqrt{(x_F - x_G)^2 + (y_F - y_G)^2 + (z_F - z_G)^2} \tag{7}$$

The tests were carried out with the MATLAB Levenberg-Marquardt algorithm using the default settings (Table. 4).

Table 4: Default MATLAB 'Levenberg Marquardt algorithm' parameters

	Value
Maximum change in variables for finite-difference gradients	Inf
Minimum change in variables for finite-difference gradients	0
Termination tolerance on the function value	1e-6
Maximum number of function evaluations allowed	100*numberOfVariables
Maximum number of iterations allowed	400
Termination tolerance on the first-order optimality	1e-4
Termination tolerance on x	1e-6
Initial value of the Levenberg-Marquardt parameter	1e-2

4.1 Results with the objective function F_1 and F_{L1}

In the following section the results of the optimization with a two dimensional F_1 and F_{L1} are presented.

Figure 6 shows the error term with different constellations of the four base stations B_i ($N = 4$). The x-axis stands for the amount of tests carried out. Every scenario was done with random constellations and random initial estimates. The y-axis represent the error term 7. The blue dots are the error with the objective function F_1 and the red dots the improved objective function F_{L1}. It can be seen that F_{L1} has no outlier. It has yet to be proven that the local minimum of F_1 becomes a saddle point for F_{L1}. However, the results show a significant effect of the F_{L1} on the optimization process.

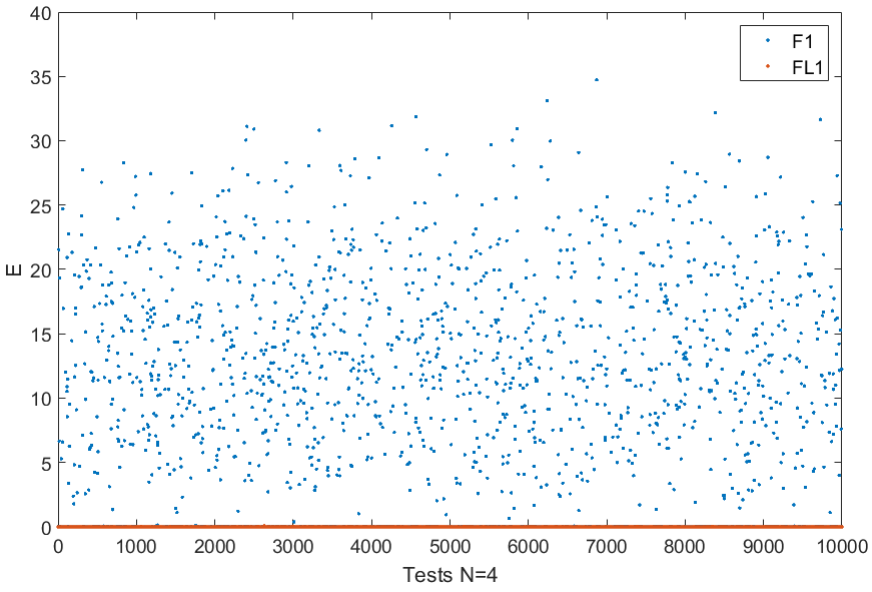

Figure 6: Blue dots: Objective function F_1. Red dots: Improved objective function F_{L1}

4.2 Results with the objective function F_2 and F_{L2}

In the following section the results of the optimization with a two dimensional F_2 and F_{L2} are presented.

Figure 7 has the same axis notations as figure 6. The x-axis stands for the amount of tests carried out and the y-axis for the error term 7. The main difference to figure 6 is that the blue dots now represent the error with objective function F_2 and the red dots the improved objective function F_{L2}. In this case it was proven that the local minimum of F_2 becomes a saddle point with an additional variable. This fact is also underpinned by the error term of F_{L2} , which is always less than 0.5.

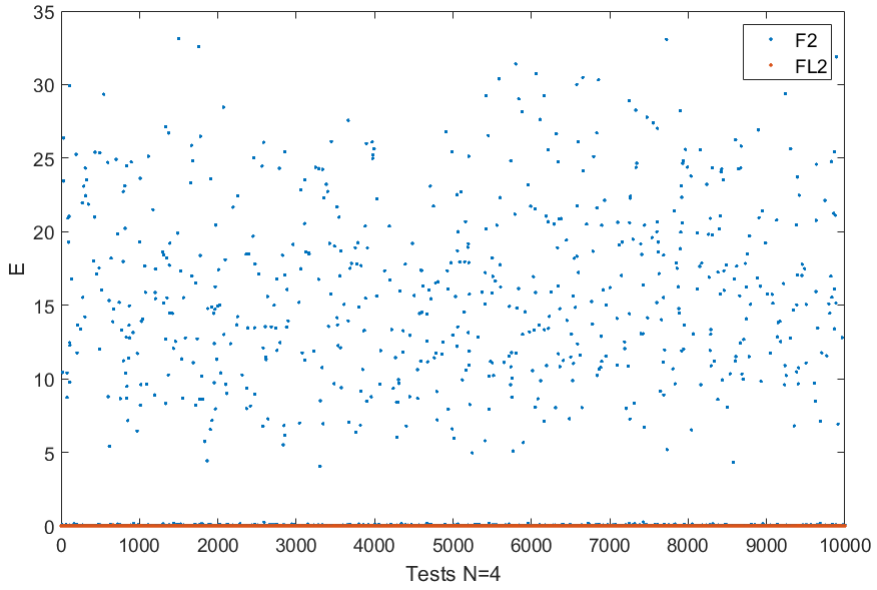

Figure 7: The blue dots are the results of the error term of F_2 . The red dots are the results of F_{L2}

4.3 Summary of the results

Tables 5 and 6 summarize the obtained results. For each number of objects (N), 10,000 constellations have been created and tested with Levenberg-Marquardt. F_L never converges to a local minimum. It can be observed that the risk to converge to a local minimum is decreasing for the objective function F_1 and F_2 with higher number of base stations. This is probable due to the increasing of the convergence radius with higher amount of base stations.

Table 5: Examples are based on a 2-D model. N: Number of base stations B_i. F_1: Objective function one, F_2: Objective function two, L: Number of local minima (Error greater than 0.5) .

N	Objective function	F_1: L	F_2: L
4	F	1357	634
4	F_L	**0**	**0**
5	F	982	399
5	F_L	**0**	**0**
6	F	810	286
6	F_L	**0**	**0**
7	F	586	182
7	F_L	**0**	**0**

Table 6: Examples are based on a 3-D model. N: Number of base stations B_i. F_1: Objective function one, F_2: Objective function two, L: Number of local minima (Error greater than 0.5)

N	Objective function	F_1: L	F_2: L
7	F	494	216
7	F_L	**0**	**0**

5 Conclusions

In the methodology section it was proven that the improved objective function F_{L2} has no local minima for non-trivial constellations. Additional to the mathematical prove a simple two dimensional example with one local minimum and one global minimum was created. This example illustrates how the optimization algorithm is using the additional dimension to bypass the local minimum and converges to the global minimum. This was underpinned by more than 100,000 numerical tests with reasonable constellations. It has still to be proven that the local minimum of F_1 becomes a saddle point with F_{L1}. However, the results show a significant effect of F_{L1} on the optimization process. The objective function F_2 performed better than the objective function F_1. Furthermore, the number of false results L decreases with a higher number of base stations B_i.

6 Discussion

It is important that the initial estimate of the additional variable is not equal to zero. Otherwise gradient-based optimization algorithms like Levenberg-Marquardt would not converge to the additional dimension. This is due to the fact that

the used optimization algorithm is estimating the stepsize for every dimension by the first derivative of the objective function. With $\lambda = 0$ are the derivatives for this additional dimension always zero, this leads to a stepsize equal zero. The test scenarios were carried out with at least for base stations. This is owed the circumstance that the additional dimension requires one more measurement. In all test scenarios the positions of base stations B_i were known. Under the condition that the reference stations are known it is also possible to obtain the solution analytically. In the case of unknown positions of base stations B_i and objects T_j it is not feasible anymore. At this point, our approach becomes extremely valuable.

7 Appendix

7.1 Proof of the assertion for objective function F_2

In Section 3.3, the assumptions and our assertion were introduced. In this section, the assertion will be proven for the objective function F_2 (2) . The proof of the assertion for objective function F_1 (1) has yet to be found. The empirical results show that the approach works for both objective functions. First, a new coordinate system is defined. This coordinate system is centered at the local minimum with global minimum on the positive x axis.

7.1.1 Definition of the new coordinate system

Without loss of generality, the following coordinate system can be used

$$\boxed{y_G = z_G = x_L = y_L = z_L = 0}$$

and

$$\boxed{x_G > 0.}$$

The distances between the base stations B_i and and object T are defined in 3.1. With the new coordinate system the equation becomes

$$d_i^2 = (x_G - a_i)^2 + (b_i)^2 + (c_i)^2.$$

The second objective function can also be written as

$$F_2(x, y, z) = \sum_{i=1}^{N} \varphi_i(x, y, z)^2$$

with the auxiliary function $\varphi_i(x, y, z)$

$$\varphi_i(x, y, z) := (x - a_i)^2 + (y - b_i)^2 + (z - c_i)^2 - d_i^2.$$

At the position of the local minimum L, the auxiliary function becomes

$$\varphi_i(0,0,0) = \left[{a_i}^2 + {b_i}^2 + {c_i}^2 - d_i^2\right] =$$

$$= \left[{a_i}^2 + {b_i}^2 + {c_i}^2 - (x_G - a_i)^2 - {b_i}^2 - {c_i}^2\right] =$$

$$= \left[{a_i}^2 - (x_G - a_i)^2\right] = \left[2a_i x_G - (x_G)^2\right] = x_G\left[2a_i - x_G\right]. \tag{8}$$

Therefore, the second objective function at the local minimum can be written as

$$F_2(0,0,0) = {x_G}^2 \sum_{i=1}^{N} \left[2a_i - x_G\right]^2. \tag{9}$$

In Section 3.3 the assumptions for the approach were presented.

$$\left(\frac{\partial^2}{\partial\lambda^2} F_L\right)(x_L, y_L, z_L, 0) < 0 \tag{10}$$

In the following it will be shown, that the assertion (10) is always correct for the improved objective function F_{L2} (4) . Equation. (11) and (12) are the first and second derivatives of objective function F_{L2} with respect to λ.

$$\left(\frac{\partial}{\partial\lambda} F_{L2}\right)(x,y,z,\lambda) = \sum_{i=1}^{N} \left[(x-a_i)^2 + (y-b_i)^2 + (z-c_i)^2 + \lambda^2 - d_i^2\right]\lambda \tag{11}$$

$$\left(\frac{\partial^2}{\partial\lambda^2} F_{L2}\right)(x,y,z,\lambda) = \sum_{i=1}^{N} \left[(x-a_i)^2 + (y-b_i)^2 + (z-c_i)^2 + \lambda^2 - d_i^2\right] + 2N\lambda^2 \tag{12}$$

At the local minimum $L(x_L, y_L, z_L)$

$$\left(\frac{\partial^2}{\partial\lambda^2} F_{L2}\right)(0,0,0,0) = \sum_{i=1}^{N} \left[{a_i}^2 + {b_i}^2 + {c_i}^2 - d_i^2\right] = \sum_{i}^{N} \varphi_i(0,0,0) =$$

$$= x_G \sum_{i=1}^{N} \left[2a_i - x_G\right] = 2x_G \sum_{i=1}^{N} a_i - N\,{x_G}^2.$$

We want to show that $\left(\frac{\partial^2}{\partial^2\lambda} F_{L2}\right)(x_L, y_L, z_L, 0) < 0$, hence we have to prove the inequality (13).

$$2x_G \sum_{i=1}^{N} a_i - N\,{x_G}^2 < 0$$

$$2\sum_{i=1}^{N} a_i < N\, x_G \tag{13}$$

In the next step, the condition at the local minimum is analyzed. The first derivative of objective function F_2 (2) equates to (14),

$$\left(\frac{\partial}{\partial x}F_2\right)(x,y,z) = \sum_{i=1}^{N}\left[(x-a_i)^2+(y-b_i)^2+(z-c_i)^2-d_i^2\right](x-a_i) =$$

$$= \sum_{i}^{N}\varphi_i(x,y,z)(x-a_i) \tag{14}$$

in combination with (8) the first derivative becomes (15).

$$\left(\frac{\partial}{\partial x}F_2\right)(0,0,0) = \sum_{i=1}^{N}\varphi_i(0,0,0)(-a_i) = \sum_{i=1}^{N} x_G\left[2a_i - x_G\right](-a_i) =$$

$$= \left[{x_G}^2\sum_{i=1}^{N} a_i - 2x_G\sum_{i=1}^{N}{a_i}^2\right]. \tag{15}$$

At the local minimum $L(x_L, y_L, z_L)$ the first derivative of objective function F_2 is equal to zero.

$${x_G}^2\sum_{i=1}^{N} a_i - 2x_G\sum_{i=1}^{N}{a_i}^2 = 0$$

$$x_G\sum_{i=1}^{N} a_i = 2\sum_{i=1}^{N}{a_i}^2 \tag{16}$$

This leads to $\sum_{i=1}^{N} a_i > 0$. The objective function F_2 has a higher result at the local minimum compared to the global minimum. It is assumed that the objective functions have no errors, therefore the result of F_2 at the global minimum must be zero.

$$F_2(0,0,0) > F_2(x_G,0,0) = 0 \tag{17}$$

$${x_G}^2\sum_{i=1}^{N}(2a_i - x_G)^2 > 0 \tag{18}$$

$$\sum_{i=1}^{N}(2a_i - x_G)^2 > 0$$

$$4\sum_{i=1}^{N} {a_i}^2 - 4x_G \sum_{i=1}^{N} a_i + N\,{x_G}^2 > 0$$

$$4x_G \sum_{i=1}^{N} a_i < 4\sum_{i=1}^{N} {a_i}^2 + N\,{x_G}^2 \tag{19}$$

The term $F_2(0,0,0)$ of (17) is replaced by (9). Equation (18) can be converted to (19). Combined with (16) the new inequality equates to (20) .

$$8\sum_{i=1}^{N} {a_i}^2 < 4\sum_{i=1}^{N} {a_i}^2 + N\,{x_G}^2$$

$$4\sum_{i=1}^{N} {a_i}^2 < N\,{x_G}^2 \tag{20}$$

7.1.2 Proof by Cauchy-Bunyakovsky-Schwarz inequality

The final step of the proof for the assertion, requires the Cauchy-Bunyakovsky-Schwarz inequality [35] for $\mathbb{R}^N$.

Here it is desired to prove that: $\quad 2\sum_{i=1}^{N} a_i < N\,x_G$

The Cauchy-Bunyakovsky-Schwarz inequality states that $|\langle \vec{x}, \vec{y} \rangle| \leq \|\vec{x}\| \cdot \|\vec{y}\|$. In our case the vectors are:

$$\vec{x} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \text{ and } \vec{y} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

Therefore, the left term $2\sum_{i=1}^{N} a_i$ of (13) must be less than or equal to $2\sqrt{N}\sqrt{\sum_{i=1}^{N} {a_i}^2}$.

$$2\sum_{i=1}^{N} a_i \leq 2\sqrt{N}\sqrt{\sum_{i=1}^{N} {a_i}^2} \tag{21}$$

From (20) it is known that $\sum_{i=1}^{N} (a_i)^2 < \frac{1}{4} N \cdot (x_G)^2$, therefore the right side of the (21) can be written as $2\sqrt{N}\sqrt{\frac{1}{4} N\,{x_G}^2}$.

The inequality becomes

$$2\sum_{i=1}^{N} a_i < 2\sqrt{N}\sqrt{\frac{1}{4} N\,{x_G}^2} = N\,x_G.$$

7.2 Possible constellations for the example

The coordinate system was described in Section 7.1.1. We want to find base station constellations with a local minimum at x_L,y_L,z_L and a global minimum at x_G,y_G,z_G .

7.3 Analysis of the first derivative

The first derivative of objective function F_2 has to be zero at the local minimum, this means

$$\left(\frac{\partial}{\partial x}F_2\right)(0,0) = \sum_{i=1}^{N}\left[({a_i}^2 - (x_G - a_i)^2)\cdot(-a_i)\right] = 0 \tag{22}$$

$$\left(\frac{\partial}{\partial y}F_2\right)(0,0) = \sum_{i=1}^{N}\left[({a_i}^2 - (x_G - a_i)^2)\cdot(-b_i)\right] = 0 \tag{23}$$

The simpelst constellation which fulfills (22) and (23) is

$$\begin{cases} \left[({a_i}^2 - (x_G - a_i)^2)\cdot(-a_i)\right] = 0 & 1 \le i \le N \\ \left[({a_i}^2 - (x_G - a_i)^2)\cdot(-b_i)\right] = 0 & 1 \le i \le N. \end{cases}$$

There are two obvious possibilities, which fulfill these equations. The first one is $a_i = b_i = 0$. The second one $a_i = \frac{1}{2}x_G$ with any b_i. The number of base stations using the first and second possibility we denote S_1 and S_2 respectively. Only sensible constellations are considered, therefore S_1 can only be one or zero.

7.4 Analysis of the second derivative

The second derivative of objective function F_2 (2) must be positive at the local minimum. This means,

$$\left(\frac{\partial^2}{\partial x^2}F_2\right)(0,0) = \sum_{i=1}^{N}\left[2a_i^2 + a_i^2 - (x_G - a_i)^2\right] > 0 \tag{24}$$

and

$$\left(\frac{\partial^2}{\partial y^2}F_2\right)(0,0) = \sum_{i=1}^{N}\left[2b_i^2 - x_G^2 + 2a_i x_G\right] > 0. \tag{25}$$

Inserting (22) into (24) leads to the first condition.

$$\sum_{i=1}^{N} 3a_i > N x_G. \tag{26}$$

The second and third condition is obtained by inserting the possibilities S_1 and S_2 into 24 and 25 respectively. Therefore, the second derivative of objective function F_2 becomes

$$\left(\frac{\partial^2}{\partial x^2}F_2\right)(0,0) = -S_1 x_G^2 + \frac{1}{2}S_2 x_G^2 > 0$$

and

$$\left(\frac{\partial^2}{\partial y^2}F_2\right)(0,0) = -S_1 x_G^2 + 2\sum_{i=1}^{S_2} b_i^2 > 0.$$

All the conditions required a local minimum at $L(0,0)$ are listed in Table 7.

Table 7: Conditions required for a local minimum at $L(0,0)$

	Conditions
1	$3\sum_{i=1}^{N} a_i > N x_G$
2	$0.5 \cdot S_2 > S_1$
3	$2\sum_{i=1}^{S_2} b_i^2 > x_G^2 S_1$

7.5 Constellations used in the example

The assumptions used in the example and the coordinates of the base stations B_i can be found in Table 8 and Table 9.

Table 8: Assumptions used for the example

Assumptions
$S_1 = 1$
$S_2 = 3$

Table 9: Coordinates of object B

Base stations $B_i$with index	X-Axis	Y-Axis
1	0	0
2	$0.5 \cdot x_G$	-2
3	$0.5 \cdot x_G$	1
4	$0.5 \cdot x_G$	3

Acknowledgments

The first author would like to thank Sebastian Bullinger, Gregor Stachowiak, Sebastian Tome for their inspiring discussion and the Fraunhofer IOSB for making this work possible. We also like to thank the reviewers and the Editor for their helpful and constructive comments that greatly contributed to improving the paper.

References

[1] David L. Adamy. *EW 102: A Second Course in Electronic Warfare*. Artech House, Boston London, 2004.

[2] Don Douglass. GPS instant navigation : A practical guide from basics to advanced techniques by kevin monahan. *Fine Edge Productions*, 1998.

[3] P. P. Bogdanov, A. V. Druzhin, A. E. Tiuliakov, and A. Y. Feoktistov. GLONASS time and UTC(SU). In *2014 XXXIth URSI General Assembly and Scientific Symposium (URSI GASS)*, pages 1–3, Aug 2014.

[4] P. Benevides, G. Nico, J. Catalão, and P. M. A. Miranda. Analysis of galileo and GPS integration for GNSS tomography. *IEEE Transactions on Geoscience and Remote Sensing*, 55(4):1936–1943, April 2017.

[5] V. Nambiar et al. SDR based indoor localization using ambient WiFi and GSM signals. In *2017 International Conference on Computing, Networking and Communications (ICNC)*, pages 952–957, Jan 2017.

[6] A. Marquez, B. Tank, S. K. Meghani, S. Ahmed, and K. Tepe. Accurate UWB and IMU based indoor localization for autonomous robots. In *2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–4, April 2017.

[7] M. Vossiek, R. Roskosch, and P. Heide. Precise 3-d object position tracking using FMCW radar. In *1999 29th European Microwave Conference*, volume 1, pages 234–237, Oct 1999.

[8] T. Akiyama, M. Sugimoto, and H. Hashizume. Light-synchronized acoustic ToA measurement system for mobile smart nodes. In *2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 749–752, Oct 2014.

[9] Juri Sidorenko et. al. Improved linear direct solution for asynchronous radio network localization (RNL). In *2017 Pacific Positioning, Navigation and Timing technology (PNT)*, 2017.

[10] J. Sidorenko, N. Scherer-Negenborn, M. Arens, and E. Michaelsen. Multilateration of the local position measurement. In *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pages 1–8, Oct 2016.

[11] H. Hmam. Quadratic optimisation with one quadratic equality constraint. *Electronic Warfare and Radar Division*, 2010.

[12] K. Batstone, M. Oskarsson, and K. Åström. Robust time-of-arrival self calibration with missing data and outliers. In *2016 24th European Signal Processing Conference (EUSIPCO)*, pages 2370–2374, Aug 2016.

[13] S. T. Birchfield and A. Subramanya. Microphone array position calibration by basis-point classical multidimensional scaling. *IEEE Transactions on Speech and Audio Processing*, 13(5):1025–1034, Sept 2005.

[14] V. C. Raykar, I. V. Kozintsev, and R. Lienhart. Position calibration of microphones and loudspeakers in distributed computing platforms. *IEEE Transactions on Speech and Audio Processing*, 13(1):70–83, Jan 2005.

[15] J. Prieto, A. Bahillo, S. Mazuelas, P. Fernández, R. M. Lorenzo, and E. J. Abril. Self-calibration of toa/distance relationship for wireless localization in harsh environments. In *2012 IEEE International Conference on Communications (ICC)*, pages 571–575, June 2012.

[16] Z. W. Mekonnen and A. Wittneben. Self-calibration method for toa based localization systems with generic synchronization requirement. In *2015 IEEE International Conference on Communications (ICC)*, pages 4618–4623, June 2015.

[17] Pratik Biswas, Tzu-Chen Lian, Ta-Chung Wang, and Yinyu Ye. Semidefinite programming based algorithms for sensor network localization. *ACM Trans. Sen. Netw.*, 2(2):188–220, May 2006.

[18] Emmanuel J. Candès, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis. *J. ACM*, 58(3):11:1–11:37, June 2011.

[19] R. Garg, A. Roussos, and L. Agapito. Dense variational reconstruction of non-rigid surfaces from monocular video. In *2013 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1272–1279, June 2013.

[20] Carl Olsson and Magnus Oskarsson. A convex approach to low rank matrix approximation with missing data. In *Proceedings of the 16th Scandinavian Conference on Image Analysis*, SCIA '09, pages 301–309, Berlin, Heidelberg, 2009. Springer-Verlag.

[21] H. Stewénius. *Gröbner Basis Methods for Minimal Problems in Computer-Vision*. PhD thesis, APR, 2005.

[22] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström. A complete characterization and solution to the microphone position self-calibration problem. In *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3875–3879, May 2013.

[23] M. Pollefeys and D. Nister. Direct computation of sound and microphone locations from time-difference-of-arrival data. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2445–2448, March 2008.

[24] M. Crocco, A. Del Bue, and V. Murino. A bilinear approach to the position self-calibration of multiple sensors. *IEEE Transactions on Signal Processing*, 60(2):660–673, Feb 2012.

[25] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino. A closed form solution to the microphone position self-calibration problem. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2597–2600, March 2012.

[26] Yubin Kuang, Erik Ask, Simon Burgess, and Kalle Åström. Understanding toa and tdoa network calibration using far field approximation as initial estimate. In *International Conference on Pattern Recognition Applications and Methods*, 2012.

[27] John S. Baras Ion Matei. Nonlinear programming methods for distributed optimization. *arXiv:1707.04598*, 2017.

[28] Ronen I. Brafman et al. Lifted optimization for relational preference rules. *ILP-MLG-SRL, Leuven, Belgium*, 2009.

[29] H.G. Bock and K.J. Plitt. A multiple shooting algorithm for direct solution of optimal control problems*. *IFAC Proceedings Volumes*, 17(2):1603 – 1608, 1984. 9th IFAC World Congress: A Bridge Between Control Science and Technology, Budapest, Hungary, 2-6 July 1984.

[30] M.R Osborne. On shooting methods for boundary value problems. *Journal of Mathematical Analysis and Applications*, 27(2):417 – 433, 1969.

[31] M. Peifer and J. Timmer. Parameter estimation in ordinary differential equations for biochemical processes using the method of multiple shooting. *IET Systems Biology*, 1(2):78–88, March 2007.

[32] J. Moré. The Levenberg-Marquardt Algorithm: Implementation and Theory. In *In Numerical analysis*, 1978.

[33] RB Thompson. Global positioning system: the mathematics of GPS receivers. *Mathematics magazine*, 1998.

[34] Stephen M Stigler. Gauss and the invention of least squares. *The Annals of Statistics*, 1981.

[35] V.I Bityutskov. Bunyakovskii inequality. *Encyclopedia of Mathematics*, 2001.