Minimal Problems and Applications in TOA and TDOA Localization

Burgess, Simon

2016

Document Version:
Publisher's PDF, also known as Version of record

Minimal Problems and Applications in TOA and TDOA Localization

Simon Burgess

Lund University

Faculty of Engineering
Centre for Mathematical Sciences
Mathematics

Mathematics
Centre for Mathematical Sciences
Lund University
Box 118
SE-221 00 Lund
Sweden

http://www.maths.lth.se/

Doctoral Theses in Mathematical Sciences 2016:7
ISSN 1404-0034

ISBN 978-91-7623-918-6 (print), 978-91-7623-919-3 (pdf)
LUTFMA-1058-2016

Printed in Sweden by Media-Tryck, Lund 2016

Contents

Abstract

The central problem of this thesis is locating several sources and simultaneously locating the positions of the sensors. The measurements captured by the sensors are time of arrival (TOA), time difference of arrival (TDOA), unsynchronized TDOA, or received signal strength indication (RSSI), all a variation of distance measurement between sensors and sources. Signals can be either sound or radio for TOA, TDOA, and unsynchronized TDOA, and radio for RSSI. To be able to simultaneously locate sensors and sources open up for many on-the-fly applications not needing a calibrated rig of sensors. By doing sensor calibration, the methods in this thesis also opens up for using much previous research in the field of TOA and TDOA localization, which has mostly dealt with locating sources from known positions of the sensors. In this thesis, several minimal problems for uncalibrated sensor network localization are studied and solved. A problem is minimal if it only needs the smallest necessary number of measurements to estimate the model parameters, thus neither making the model parameters over- nor underdetermined. Apart from revealing understanding and theoretical aspects of the problem, studying minimal problems also have interesting applications when dealing with larger measurement sets containing severe outliers. This thesis utilizes the random sample consensus method (RANSAC), that uses the minimal algorithms developed in this thesis, to do localization of the sensors and sources and simultaneously weed out outliers in the measurements. The set of inliers and parameters are then used in non-linear optimization schemes to refine the parameters. Experiments show that for experiments with sound, microphone and sound sources can be located with centimeter precision. For solving the minimal problems, techniques from linear algebra and multivariate polynomial solving are utilized. This thesis further investigates simultaneous localization of cell phone users and mapping of the radio environment in multi-floor environments, using RSSI measurements and pressure sensors. Nonlinear optimization and filtering techniques are used to do parameter estimation, and results in two buildings with several floors indicates that these methods can be deployed with errors in the range of 10-20m horizontally, with $> 95\%$ accuracy in floor detection.

Acknowledgements

I would like to thank my supervisor, Prof. Kalle Åström, for always providing me with inspiration and ample opportunities to progress our research in a direction to my liking. When speaking to my fellow Ph.D. students around the world, often they have to lighten their heart with grievances directed towards their supervisors. Every time, I have counted myself lucky to have been Kalle's Ph.D. student. Even when he was head of department, his door was always open for a chat with me. He has helped create a welcome and open work environment, where my own interests have dominated my direction. If I have any complains, it is that he is too smart; sometimes he solves an interesting problem during the bike ride home, whereas I would have liked to have a little more time.

My co-authors, many of which are my colleagues at the Centre for Mathematical Sciences, deserve a big thank you. Many stimulating conversations and learning experiences were happily provided. I would also like to especially mention my co-authors from University of Freiburg, Prof. Christian Schindelhauer and Dr. Johannes Wendeberg, and my co-authors from Combain Positioning Solutions, for collaborating when our research interests aligned.

One of my most precious experiences during my time as a Ph.D. student was teaching the introductory mathematical statistics classes. A big thanks to Lena Zetterqvist, Maria Sandsten, Andreas Jakobsson, and the staff at Mathematical Statistics for giving me that wonderful opportunity.

Thanks to my lovely friends at Skydive Skåne for helping me alleviate the stress I have felt working towards this thesis. I cannot imagine these years without you. Kenneth, I will always remember you and all the things you taught me, both about skydiving and about life. Fly free.

Finally, I would like to thank my wife Annie, and my son Walde, for putting up with me.

Lund, 2016 *Simon Burgess*

List of papers

This thesis is based on the following papers:

A Simon Burgess, Yubin Kuang, and Kalle Åström, "TOA Sensor Network Self-Calibration for Receiver and Transmitter Spaces with Difference in Dimension", *Elsevier Signal Processing*, 2015.

B Yubin Kuang, Erik Ask, Simon Burgess, and Kalle Åström, "Understanding TOA and TDOA Network Calibration using Far Field Approximation as Initial Estimate", *International Conference on Pattern Recognition Applications and Methods*, Algarve, Portugal, 2012.

C Simon Burgess, Yubin Kuang, Johannes Wendeberg, Kalle Åström, and Christian Schindelhauer, "Minimal Solvers for Unsynchronized TDOA Sensor Network Calibration using Far Field Approximation", *International Symposium on Algorithms and Experiments for Sensor Systems, Wireless Networks and Distributed Robotics*, Sophia Antipolis, France, 2013.

D Zhayida Simayijiang, Simon Burgess, Yubin Kuang, and Kalle Åström, "TOA-Based Self-Calibration of Dual-Microphone Array", *IEEE Journal on Selected Topics in Signal Processing*, 2015.

E Yubin Kuang, Simon Burgess, Anna Torstensson, and Kalle Åström, "A Complete Characterization and Solution to the Microphone Position Self-Calibration Problem", *International Conference on Acoustics, Speech and Signal Processing*, Vancouver, Canada, 2013.

F Simon Burgess, Mikael Högström, Björn Lindquist, and Kalle Åström, "Smartphone Positioning in Multi-Floor Environments Without Calibration or Added Infrastructure", to appear in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Madrid, Spain, 2016.

Additional papers not included in the thesis:

1. Stefan Ingi Adalbjörnsson, Ted Kronvall, Simon Burgess, Kalle Åström, and Andreas Jakobsson, "Sparse Localization of Harmonic Audio Sources", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2016.

2. Zhayida Simayijiang, Simon Burgess, Yubin Kuang, and Kalle Åström, "Mininal Solutions for Dual Microphone Rig Self-Calibration", *European Signal Processing Conference*, Lisbon , Portugal, 2014.

3. Erik Ask, Simon Burgess, and Kalle Åström, "Minimal Structure and Motion Problems for TOA and TDOA Measurements with Collinearity Constraints", *International Conference on Pattern Recognition Applications and Methods*, Barcelona, Spain, 2013.

4. Simon Burgess, Yubin Kuang, and Kalle Åström, "TOA Sensor Network Calibration for Receiver and Transmitter Spaces with Difference in Dimension", *European Signal Processing Conference*, Marrakesh, Morocco, 2013.

5. Simon Burgess, Yubin Kuang, and Kalle Åström, "Node Localization in Unsynchronized Time of Arrival Sensor Networks", *International Conference on Pattern Recognition*, Tokyo, Japan, 2012.

6. Simon Burgess, Yubin Kuang, and Kalle Åström, "Pose Estimation from Minimal Dual-Receiver Configurations", *International Conference on Pattern Recognition*, Tokyo, Japan, 2012.

Introduction

This thesis is focused on algorithms used in simultaneously locating events and calibrating sensor arrays using time of arrival (TOA) measurements, time difference of arrival (TDOA) measurements, or derivations of these measurements. The classical applications are in localization in radio and sound applications, which this work provides several examples of. The focus is on minimal algorithms, i.e. algorithms that use no more than the necessary number of receivers in the sensor array, and no more than the necessary number of transmitting events. Apart from often revealing understanding and theoretical aspects of the problem, minimal problems also have interesting applications when dealing with larger measurement sets containing severe outliers. Considerations are given to expand and deploy the minimal sensor network calibration algorithms to deal with these bigger measurement sets.

In this section, some of the underlying theory used in this thesis is introduced, as well as a background of the active field of localization where this work fits in.

1 Background

Positioning and navigation has throughout the ages been a key instinct for survival in the natural world. Today, the need for navigation is still an integral part of modern life, present everywhere from finding your keys, to first responders being able to locate where an emergency call was made from with accuracy. Positioning of e.g. artillery using sound has been done in theory since at least 1741, [1], although widespread applications and interest first became prevalent during the 20th century. In World War I the position of enemy artillery pieces was located using passive microphones at predetermined positions, [2]. Models for sound localization from a biological perspective were also begun to be studied during the 20th century, [3].

The Global Navigation Satellite System (GNSS) emerged at the end of the 20th century from military applications, but soon found its way into many civilian applications. The GNSS receivers use either the GPS, GLONASS, Galileo or

BeiDou system to fix an absolute position, and applications are widespread. The basics of positioning is trilateration using time-of-flight from three or more time synchronized satellites, [4]. Root Mean Square Errors (RMSE) can be as low as a couple of centimeters, with an initialization time of < 10s in good conditions. However, in practice, the RMSE is often a couple of meters, and furthermore works poorly or not at all indoors.

For indoor radio based localization, [5] provides an overview. The approaches can be divided into four categories: i) Proximity based, which uses a binary "there or not" approach. These are fast, easy to implement, but imprecise. ii) Trilateration-like based. Here we find TOA and TDOA localization, as well as methods working with Received Signal Strength (RSS). iii) Triangulation based, or Angle of Arrival (AOA). iv) Pattern recognition based, or fingerprinting. These use a previously collected database of measurements together with ground truth. Any new measurement that needs to be positioned needs to have its position derived from the database measurements.

All of these generally use prior calibration in the form of deployment of specialized hardware annotated with ground truth, or extensive ground truth data collection. There have been several attempts to alleviate this problem. We denote this calibration-free, or uncalibrated, localization. In [6] and later [7], the Cramér-Rao Lower Bound (CRLB) is derived for systems doing calibration-free localization, using RSS, TOA or AOA signals between nodes. Noise is assumed to be Gaussian on TOA and AOA measurements, and log-normal on RSS measurements. Some of the nodes must have known locations, so-called anchors.

There are several systems using WiFi calibration-free localization that use extra sensors or other external modalities to help. In [8], good results around 2m accuracy are obtained, but the method also uses Inertial-Measurement Units (IMU) for pedometry and gyroscopes. In [9–12], systems are deployed with good results, but rely heavily on a floor map. Furthermore, IMUs for pedometry, magnetometers, and/or gyroscope measurements are used to do the calibration-free localization.

As seen, many of the approaches for indoor radio based localization and calibration use sensor fusion. In this thesis it is explored what can be done using only TOA, TDOA, or RSS, and look at the theoretical minimal number receivers and transmitters needed. Besides giving a deeper understanding of each modality, and what is needed to do calibration, it is argued that this opens up for many on-the-fly applications, and alleviates the need for specialized hardware. Furthermore,

TOA and TDOA calibration-free localization can be applied to sound, and thus any source of sound becomes a transmitter.

For sound based localization, most previous work has been focused on how to locate one or more sources or source directions from a calibrated array of microphones. A survey of the methods can be found in [13]. The problem of locating one or several sound sources using a calibrated array of microphones also has interesting applications, like tracking a sound source in real-time, or boosting a specific source, [14]. It is not without challenges though, as (i) sound sources might not be constantly emitting sounds, and the number of sources may change over time, (ii) indoor environments are often reverberant, (iii) the microphone rig often brings its own difficulties in capturing and modeling the signal, as the microphones are often embedded in specific hardware, e.g. a robot or a special non-uniform microphone array. For localization, there are two different approaches often used. The first tries to calculate the time delays between microphones for interesting sound events, and then use these time differences to do multilateration, cf. [15–18]. Common techniques for delay estimation include different variations on cross-correlation or canonical correlation analysis, which then allows the sources to be located in a second step using multilateration. A popular one is Generalized Cross Correlation with Phase Transform (GCC-PHAT), presented in 1976, [19]. These methods are often fast, but often deteriorate fast in environments with noise or reverberation, due to many of the calculated time differences being outliers.

The second strategy takes a one-step approach, where TDOA measurements and positions are considered more or less simultaneously with the sound signals as input. An example is the steered response power phase transform (SRP-PHAT) method, [20], that has for long time been a reference for real time sound source tracking and localization, and has seen many successors, e.g. [21, 22].

Other uses for calibrated microphone arrays are to use echoes to reveal room shapes, from known sounds and sound sources, [23]. Common for all of these methods is the need for a calibrated rig of microphones. On-the-fly calibration could again open up for many interesting applications without the need for a pre-calibrated rig. For an overview of automatic microphone array calibration techniques, spanning over several different types of position-related measurements used, and the different necessities of synchronization, see [24].

2 Fundamentals of TOA/TDOA localization

Here follows an overview of the fundamentals of TOA and TDOA localization. The first part deals with the calibrated case, and the second part gives an overview of the uncalibrated case and relevant prior work. Consider the case where the receivers are calibrated, e.g. have known positions, and we have omnidirectional sensors. For simplicity, we consider the signals to be sounds in this section, but the theory applies to radio signals as well. We want to locate an incoming sound event with index j, which has an emanating position $\mathbf{s}_j$, using the time stamps of the event arriving at the sensors with index i and positions $\mathbf{r}_i$. The positions $\mathbf{s}_j$ and $\mathbf{r}_i$ can be in either $\mathbb{R}^2$, i.e 2D, or $\mathbb{R}^3$ i.e. 3D. Higher dimensional cases can be of theoretical interest, as well as different fields like the algebraically complete $\mathbb{C}^d$, and are indeed considered in many of the papers in this thesis, but here we stick with 2D and 3D. It is here also assumed that i) the speed of the medium, v, is known and constant, ii) the times of a sound event impinging on each sensor can be measured, and iii) different events can be identified by the sensors, i.e. the event j impinging on a sensor can be identified as coming from event j. In practice, iii) can be done with events being separated in time, by using different frequencies, or matching algorithms, cf. [13]. This is however not perfect, and how to cope with poor or errenous measurements is briefly addressed in Section 3.

For the TOA scenario, the measurement is simply the time, t_{ij}, it takes for the signal to travel from event j to sensor i. For the TDOA scenario, the time when the event occurred originally at its source is unknown, and thus the measurement becomes $t_{ij} + o_j$, where o_j is the time when the event occurred. By multiplying the measurements with the known speed of the medium, v, we get the measurements in distances,

$$\begin{aligned} d_{ij,TOA} &= |\mathbf{r}_i - \mathbf{s}_j|, \\ d_{ij,TDOA} &= |\mathbf{r}_i - \mathbf{s}_j| + f_j. \end{aligned} \tag{1}$$

In the TOA case, it is simply the distance between sensor i and event position j. In the TDOA case, the time o_j when the event occurred has become a distance offset f_i. The physical interpretation for the offset constant may not be as clear, but as we generally are interested in the positions, distances are more convenient to work with. By normalizing the measurements as $\bar{d}_{ij} = d_{ij} - d_{1j}$, the new constant depending on j will be $\bar{f}_j = |\mathbf{r}_1 - \mathbf{s}_j|$, which has a clearer interpretation.

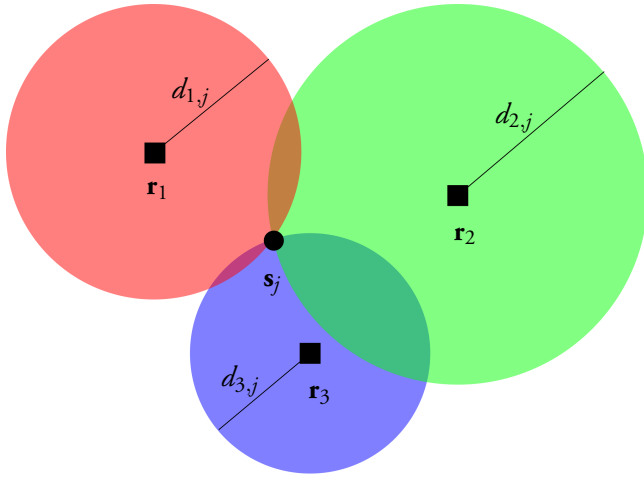


Figure 1: TOA localization with known sensor positions $\mathbf{r}_i$, shown as squares. The distance measurements d_{ij} tells us how far from $\mathbf{r}_i$ the event should be. The black circle is the event position to be located, $\mathbf{s}_j$.

For TOA localization, each measurement $d_{ij,TOA}$ trivially gives a solution space for the position of the event $\mathbf{s}_j$, as a sphere around $\mathbf{r}_i$. Several measurements give the solution space as the intersection of all spheres, or for an overdetermined solution, the point that in some sense is closest to all spheres. See Figure 1 for a visualization.

For TDOA localization, each separate measurement $d_{ij,TDOA}$ gives no information of the position $\mathbf{s}_j$ of the event, because of the additive constant f_j in (1). But the difference between two measurements from different sensors coming from the same time event does,

$$d_{ij,TDOA} - d_{kj,TDOA} = |\mathbf{r}_i - \mathbf{s}_j| - |\mathbf{r}_k - \mathbf{s}_j|. \tag{2}$$

The difference in the measurements is the difference in distance for the event to reach sensors i and k, and is what gives rise to the name Time *Difference* Of Arrival. One such difference gives that the position of the emanating event $\mathbf{s}_j$ lies on one half of a hyperbola if in 2D, or one half of hyperboloid of two sheets in 3D. These geometrical spaces can be defined from the exact property we are looking for, that the difference in distance to two points, in this case $\mathbf{r}_1$ and $\mathbf{r}_2$, is

constant. For more than two sensors, each difference of measurements like in (2) that is not a linear combination on any other differences of measurements gives rise to a new half hyperbola (or half hyperboloid of two sheets in 3D) that restricts the solution space. Thus, for n sensors giving d TDOA measurements, we get the $n-1$ hyperbolas restricting the solution space for each pair of receivers. The solution, i.e. position $\mathbf{s}_j$, lies on the intersection of these half hyperbolas.

In the case where we have more measurements than unknowns, and the measurements are affected by noise, the solution spaces of half hyperbolas generally do not all intersect. The best solution then becomes a function of the noise model for the measurements. Perhaps some should be removed completely due to being outliers, and the rest assumed to follow some noise model. See Section 3 for more examples of this.

Many times, the TDOA is measured with the phase shift of the incoming event. Assume that the signal has a wave length of λ. The two sensors receive the signal, but detect a phase shift $-\pi/2 \leq \alpha \leq \pi/2$ between the received signal at sensor 1 and sensor 2. If the distance between two receivers is $< \lambda/2$, the phase shift can trivially be translated as a distance difference, and we are back to the situation in (2) and the same solution space as described in previous paragraphs. But if the distance between the two sensors is $\geq \lambda/2$, depending on the phase shift α, several possible distance differences between the event and sensor 1, and event and sensor 2, are possible. The possible distance differences Δ creating the measured phase shift α are the correct distance difference in (2), $\pm\lambda/2$ such that the possible distance difference is $\leq$ the total distance between the two sensors, $||\mathbf{r}_i - \mathbf{r}_k||$. The possible distance differences be written as

$$\Delta = \left\{ \delta = |\mathbf{r}_i - \mathbf{s}_j| - |\mathbf{r}_k - \mathbf{s}_j| + c\lambda/2 : c \in \mathbb{Z}, \delta \leq |\mathbf{r}_i - \mathbf{r}_k| \right\}. \tag{3}$$

Each such possible distance difference in Δ gives rise to one half of a hyperbola in 2D, or half hyperboloid of two sheets in 3D. So for each pair of sensors where at least one of the sensors has not already been used in a pair, we now get not only one surface, but possibly several surfaces with feasible positions for $\mathbf{s}_j$. The solution lies on the intersection of the surfaces coming from each pair of sensors. See Figure 2 for a visualization.

When the event position is very far from the sensors, the incoming wave impinging on the sensors is almost flat. The problem of determining the distance to the event position then becomes ill-conditioned, as the wave front does not

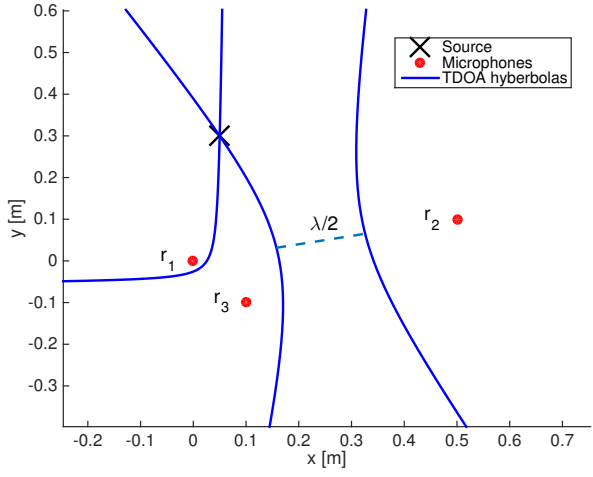


Figure 2: TDOA hyperbolas representing all feasible locations of a single source received by three sensors. As $|\mathbf{r}_2 - \mathbf{r}_1| > \lambda/2$, spatial aliasing yields another hyperbola of feasible locations. And yet, in this case, there exists only one intersection between the hyperbolas obtained from different pairs, and so the estimate of the source may still be obtained unambiguously.

change much if you move the event position a bit towards or away from the sensors. However, determining the direction to the event position can still be done. Thus, TOA/TDOA localization also requires algorithms relying on the far-field assumption, e.g. that the sound event is far away from the sensors and thus can be said to have one common direction to the sensor array.

2.1 Uncalibrated TOA/TDOA localization

The case where the sensor positions are not known is referred to as uncalibrated TOA/TDOA localization, or simply TOA/TDOA calibration. In the process of recovering both the sensor positions and event positions calibration occurs, as both $\mathbf{s}_j$ and $\mathbf{r}_i$ are inferred from the measurements in (1).

Without any other measurements to anchor the system, a solution of positions $\mathbf{s}_j$ and $\mathbf{r}_i$ can always be rotated, translated and/or mirrored while still fulfilling the measurements in (1), due to the measurements only using the positions

for relative distances. These ambiguities in solutions, and other degrees of freedom that cannot be uniquely determined by more measurements, are referred to as gauge freedom. The more common setting for uncalibrated localization is TDOA, where the time a source signal is emitted is unknown. But the case of array calibration when TOA measurements are available is also of importance in the corresponding TDOA calibration, where a stratified approach of first determining offsets f_j in (1) and then solving the TOA calibration problem has been used, see e.g. [25, 26].

It is often harder to infer geometrical interpretations for the solutions, compared to the calibrated case. In paper C, some geometrical properties are derived for TDOA calibration in the case when the event positions are far away from the sensors, i.e. a far-field setting. The differences (2) of sensors 2,3 and 4 as k, and sensor 1 as i should in the noiseless case lie on an ellipsoid in 3D, and an ellipse in 2D.

Some previous work has been done in uncalibrated TOA/TDOA localization. Several previous contributions dealing with sensor network calibration rely on prior knowledge or extra assumptions about locations of the sensors to initialize the problem, see [27–32]. Being able to do uncalibrated localization without these extra assumptions opens up for a wider range of applications, and papers A-E deal with cases of TOA/TDOA calibration without these assumptions.

In [33] a far-field approximation was utilized to solve the TOA and TDOA case for 2D. TOA calibration using only measurements has been studied in [34], where a closed-form solution to the minimal case of three transmitters and three receivers in the plane are given. In general, there are three solutions for a given set of measurements. Calibration of TDOA networks was studied in [25] and further improved upon in [26], and provides closed-form solvers although the minimal cases are still unsolved. In [35, 36], a TDOA setup is used for indoor navigation based on non-linear optimization, but the methods can get stuck in local minima and are dependent on initialization.

The problem relates to the study of sensor networks under rigid graph theory [37, 38] where general graph structure is of interest. The TOA self-calibration problem studied here corresponds to a special case - bipartite graphs [39].

3 Minimal problems and RANSAC

One can ask how few measurements, (1), between sensors and events are needed to be able to get a finite number of solutions to the positions of sensors and events. Besides revealing interesting theoretical aspects of the problem at hand, determining and solving the problem with a minimal number of measurements has intriguing applications. The idea is that using a few number of measurements has a higher probability of not using any outliers. This thesis will frequently deal with constructing and applying such algorithms. Using only the necessary number of measurements will be referred to as minimal problems, or minimal cases.

Minimal problems have a long history in a wide area of applications. In computer vision, the problem of estimating the fundamental matrix for stereo images using uncalibrated cameras requires seven point correspondences between the images, and has one or three real solutions. This was studied as early as 1855, [40].

Minimal problems have been studied extensively for computer vision and image analysis applications, [41], where measurements often are quite accurate, or outliers. This setup is well suited for the random sample consensus (RANSAC) algorithm, [42], which simultaneously estimates parameters for a model and identifies a set of inliers. Although RANSAC has primarily been applied in computer vision and image analysis, it has been used by the signal processing community for calibrated audio localization, [14, 43, 44].

Briefly, RANSAC works as follows. Select randomly a minimal or close to minimal number of measurements from the data set and fit a hypothesized model using a minimal algorithm to the selected measurements. Count how many of the total number of measurements that fit with the hypothesized model using a threshold to see if it fits or not. If all points selected for model fitting were inliers, there should be other measurements that are close to the model. Repeat the random minimal sampling and parameter estimation until a big enough inlier set has been found, and return the corresponding parameter and inlier set hypothesis.

The RANSAC method can be illustrated with an example - here we study the problem of finding a circle with fixed radius $r = 0.2$. For this problem there are two unknown parameters, the x and y coordinates of the circle center (u, v). The measurement data $(x_1, y_1), \ldots, (x_n, y_n)$ contains both inliers, points for which

$$(x_i - u)^2 + (y_i - v)^2 \approx r^2, \tag{4}$$

and outliers, i.e. points for which

$$|(x_i - u)^2 + (y_i - v)^2 - r^2| \gg 0. \tag{5}$$

Such a point set is illustrated in Figure 3. Notice the distribution of inliers and outliers. For such a point set, a standard least squares or L_1 optimization typically fails at providing good estimates of the parameters.

First a *hypothesize step*, in which a minimal subset of the data is randomly selected. In this case the minimal problem is to solve for the circle center given two random points. This sub-problem typically has two solutions, i.e. there are two circles that go through the selected points. For each of the solutions, one calculates the number of inliers in the *test step*. For each measurement point one calculates the residual,

$$r_i(u, v) = (x_i - u)^2 + (y_i - v)^2 - r^2, \tag{6}$$

and checks the number of inliers, i.e. those that have $|r_i(u, v)| < T$, where T is a pre-defined threshold. This process of hypothesize and test is repeated a fixed number of iterations. The parameter (u, v) that gave the highest number of inliers is chosen as initial estimate of the parameter estimation problem.

Two hypothesize and test iterations of the algorithm are illustrated in Figure 3. In the figure, measurement points $(x_1, y_1), \ldots, (x_n, y_n)$ containing both inliers and outliers to the circle fitting problem are shown as dots. In one of the iterations of the RANSAC algorithm, two random points, shown as red stars, are selected. This subset of data provides two hypothetical circles, for which the inlier count (the number of points between the two red circles) is low. In the plot the circles for $|r_i(u, v)| = T$ are shown as solid red lines for one of the two solutions and as dashed lines for the other solution. For another iteration two random points (shown as green stars) provide two solutions to (u, v). One of them gives a large number of inliers.

An advantage to RANSAC is that for fast minimal solvers that are at the core, for many applications one can efficiently get a good estimate of the model and inlier set. A disadvantage to RANSAC is that there is no guarantee of optimality, but rather a probabilistic reasoning that given enough iterations, a good minimal inlier set will eventually be selected in a hypothesize step, giving decent parameter estimates and inlier identification. Another disadvantage is that the score of how good a parameter hypothesis is, is the cardinality of the inlier set, defined by the

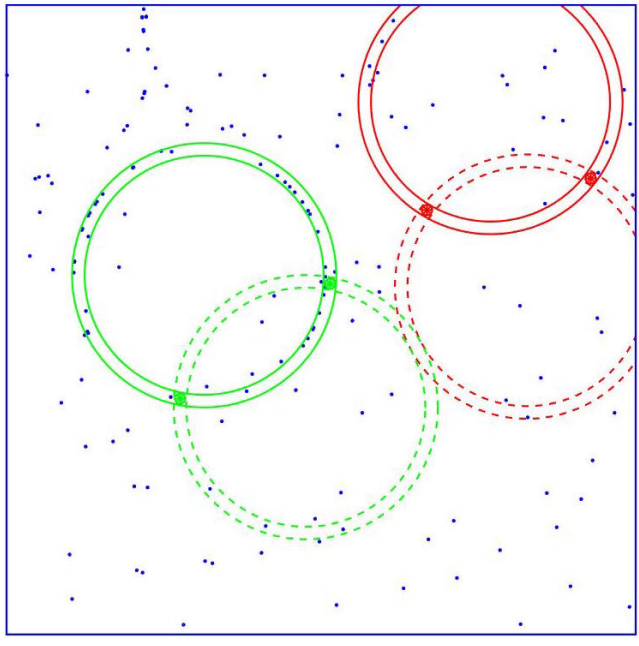


Figure 3: The figure illustrates measurement points $(x_1, y_1), \ldots, (x_n, y_n)$ containing both inliers and outliers to the circle fitting problem.

threshold. A large threshold will tend to make all hypotheses equally good, and a small threshold will tend to make the estimated parameters unstable.

Several strategies have been used to try to remedy these shortcomings. In [45, 46], a strategy for finding the optimal set of inliers using the thresholding criterion is implemented. In [47], the prior probabilities of a measurement being an inlier is used in computing the inlier set, as well as the quality of the inlier set being a maximal likelihood estimation, instead of just the cardinality. In general though, RANSAC is widely used within image analysis and computer vision as is, [41].

Many of the minimal problems in this thesis involve the measurements in (1), and by squaring the equations, we get

$$d_{ij,TOA}^2 = |\mathbf{r}_i - \mathbf{s}_j|^2, \tag{7}$$

$$\left(d_{ij,TDOA} - f_j\right)^2 = |\mathbf{r}_i - \mathbf{s}_j|^2. \tag{8}$$

These equations are now multivariate polynomials in the unknowns in the vectors $\mathbf{r}_i$, $\mathbf{s}_j$, and f_j. In general, solving for the unknowns up to gauge freedom by directly applying polynomial solvers to these equations is out of reach for state-of-the-art polynomial solvers. For instance, in the case of solving the TOA calibration problem in 3D, presented in paper E, in general there needs to be either 4 receivers and 6 transmitters, or 6 receivers and 4 transmitters, thus including 30 unknowns. There are three unknowns for the rotation of a solution that can never be fixed, and 3 for the translation of a solution, so after accounting for that by for example fixing some of the unknowns, there are still 24 unknowns. This is in general far too many unknowns for quadratic equations for polynomial solvers to handle. Thus, the problem needs to be addressed in a different manner. Nevertheless, often these problems can be reformulated and worked until they can be solved for fewer unknowns in a polynomial system of equations.

4 Solving polynomial equations with the action matrix method

This section presents some fundamentals of the theory of algebraic geometry, the field of studying and solving systems of polynomial equations over $\mathbb{C}$, or in general any algebraically closed field. The action matrix method is further presented, which is used in several papers of this thesis, and some recent advances of the method. See [48] for a thorough explanation of the claims made here.

The problem at hand is to find solutions to

$$\begin{aligned} f_1(\mathbf{x}) &= 0, \\ f_2(\mathbf{x}) &= 0, \\ &\vdots \\ f_n(\mathbf{x}) &= 0, \end{aligned} \tag{9}$$

where $f_i(\mathbf{x})$ are multivariate polynomials of the unknowns $\mathbf{x} \in \mathbb{C}$.

The ring of polynomials over the field of complex numbers is usually denoted $\mathbb{C}[\mathbf{x}]$. Let V denote the solution set to (9). In algebraic geometry, V is called an algebraic variety.

Let $I = \{\sum_{i=1}^{n} h_i f_i : h_i \in \mathbb{C}[\mathbf{x}]\}$, i.e. the set of all sums of the polynomials in (9), each multiplied with any polynomial. This is called the ideal generated by $f_1, f_2, \ldots, f_n$. One aspect of why ideals are interesting to study is that they generalize the system of polynomials at hand: $\mathbf{x}$ is a solution to (9) iff it is a zero of I. One could also ask what the set of polynomials that vanish on V is. Clearly, I is contained in this set. If the opposite holds, I is called radical. For example, the polynomial equation in one variable $x^2 = 0$ does not have a radical ideal as $I = \{hx^2 : h \in \mathbb{C}[\mathbf{x}]\}$ does not contain x, a polynomial that vanishes on V. The ideal generated by $x = 0$, however, is trivially radical.

We say that two polynomials f and g are equivalent modulo I, or $f \sim g$, if $f - g \in I$. This equivalence relation gives rise to a partition of $\mathbb{C}[\mathbf{x}]$. The quotient space $\mathbb{C}[\mathbf{x}]/I$ is the set holding the partitions, or equivalence classes. Let $[\cdot]$ be the natural projection from a polynomial in $\mathbb{C}[\mathbf{x}]$ to $\mathbb{C}[\mathbf{x}]/I$, i.e $[f]$ is the equivalence class that has f in it.

One can also consider the equivalence classes of all polynomial functions that are equal on V. Denote this set of equivalence classes by $\mathbb{C}(V)$. A polynomial function on V is a function from V to $\mathbb{C}$ that can be exactly described by a polynomial at the points of V. If polynomials are equivalent modulo I, they are trivially equal on V. The converse holds if I is radical. Thus, we get to the conclusion that, for racial ideals I, $\mathbb{C}[\mathbf{x}]/I$ and $\mathbb{C}(V)$ are isomorphic. Here, we generally only consider finite V, as minimal problems generally only have a finite number of solutions. For V finite, any function on V to $\mathbb{C}$ is a polynomial function, due to the unisolvence theorem for polynomials, which says that any function on a finite set of points can be interpolated exactly by a polynomial. Thus, any equivalence class in $\mathbb{C}(V)$ can be represented by a vector in $\mathbb{C}^k$ with $k = |V|$ elements. As operations translate through, $\mathbb{C}(V)$ and $\mathbb{C}^k$ are isomorphic. The final conclusion is thus that, for finite algebraic varieties V and radical ideals I, the quotient space $\mathbb{C}[\mathbf{x}]/I$ and $\mathbb{C}^k$ are isomorphic.

4.1 The action matrix method

This section describes the fundamentals of the action matrix method (sometimes referred to as the Gröbner basis method) for solving polynomial equations. This is used for polynomial equations with a finite number of solutions, which is gen-

erally the case with minimal problems. If V is finite, $\mathbb{C}[\mathbf{x}]/I$ is finite dimensional, and if I is radical, $\mathbb{C}[\mathbf{x}]/I$ has the same dimension as $|V|$. Now consider the operator $T_a : \mathbb{C}[\mathbf{x}]/I \to \mathbb{C}[\mathbf{x}]/I$ such that $[f(\mathbf{x})] \mapsto [a(\mathbf{x})f(\mathbf{x})]$. This operation is linear, and since $\mathbb{C}[\mathbf{x}]/I$ is finite dimensional, we can select a linear basis for $\mathbb{C}[\mathbf{x}]/I$ and represent the T_a as a matrix multiplication. Denote this matrix as M_a. The eigenvalues to M_a^T are $a(\mathbf{x})$ evaluated at V, and the eigenvectors are the basis elements evaluated at V. To see this, consider any equivalence class $[r(\mathbf{x}) = c^T\mathbf{b}]$, where $\mathbf{b}$ is a vector of polynomials forming a basis for $\mathbb{C}[\mathbf{x}]/I$, and c is a vector of coefficients. Now, the operation T_a on $[r(\mathbf{x}) = c^T\mathbf{b}]$ can be written as

$$[a(\mathbf{x}) \cdot c^T\mathbf{b}] = [(M_a c)^T\mathbf{b}] = [c^T M_a^T\mathbf{b}]. \tag{10}$$

As this holds for any coefficients c, we get that $[a(\mathbf{x})\mathbf{b}] = [M_a^T\mathbf{b}]$. This means that

$$a(\mathbf{x})\mathbf{b} = M_a^T\mathbf{b} + \mathbf{g}(\mathbf{x}) \tag{11}$$

for some vector $\mathbf{g}(\mathbf{x})$ with elements $g_i(\mathbf{x}) \in I$. Evaluating this in points $\mathbf{x} \in V$, we get that

$$a(\mathbf{x})\mathbf{b} = M_a^T\mathbf{b}, \tag{12}$$

which means that the basis $\mathbf{b}$ evaluated at $\mathbf{x}$ consist of eigenvectors of M_a^T, and corresponding eigenvalues are $a(\mathbf{x})$.

It remains how to calculate a basis for $\mathbb{C}[\mathbf{x}]/I$. We need to have a well defined representative for each equivalence class $[f] \in \mathbb{C}[\mathbf{x}]/I$. Generally this is done by polynomial division over the polynomials in (9). For multivariate polynomials, however, the remainder under division by the polynomials is not unique, and depends on the order of the division. This is solved by calculating a Gröbner basis for I. A Gröbner Basis is a set of generating polynomials for I that has the property that polynomial division with the generating polynomial always has a well defined remainder. Any element $[f] \in \mathbb{C}[\mathbf{x}]/I$ can thus be identified with its remainder under division by the Gröbner basis. A Gröbner basis can be calculated in finite time with Buchberger's algorithm. Once a Gröbner basis has been calculated, the basis for $\mathbb{C}[\mathbf{x}]/I$ can for instance be chosen by the following

definition of a Gröbner basis: The leading term of any polynomial in I is divisible by the leading term of some polynomial in the Gröbner basis. The leading term here is the monomial of the highest order, according to some monomial order, with its coefficient. So if it is not in I, e.g. will have a remainder under polynomial division, the leading term will not be divisible by any leading term in the Gröbner basis of I. Then, we can use all the monomials that are not divisible by any leading monomial in G as a basis.

4.1.1 Problems and advances for the action matrix method

Buchberger's algorithm can compute a Gröbner basis for an ideal I in finite time, but is numerically unstable in floating point arithmetic, due to propagated round-off errors. In [49], emulated 128-bit precision arithmetic is used to make the Gröbner basis calculation numerically stable, but renders the solution too slow for many practical purposes. Some use a hand-tailored Gröbner basis for the specific problem at hand, [50].

Byröd et al., [51], developed solvers using the strategy that a Gröbner basis is not actually needed for constructing the action matrix. Not all equivalence classes $[f]$ need to be represented in $\mathbb{C}[\mathbf{x}]/I$, but only $[f]$ that are in $a(\mathbf{x})\mathbf{b} \setminus \mathbf{b}$, where $\mathbb{B}$ is the considered basis for $\mathbb{C}[\mathbf{x}]/I$. This give huge freedom on how to select a basis, which done right, buys numerical stability.

The general strategy used for polynomial solving in this thesis, following [51], is briefly explained below.

1. Select a monomial as multiplying polynomial, $a(\mathbf{x})$.

2. Expand the equation system in (9) by multiplying the equations with new monomials to create more equations. Formulate it as a matrix multiplication of coefficients times monomials.

3. Among all the monomials, select a suitable candidate for a basis $\mathbf{b}$ of $\mathbb{C}[\mathbf{x}]/I$. The monomials in $a(\mathbf{x})\mathbf{b} \setminus \mathbf{b}$ need to be reduced over I to be expressed in terms of $\mathbf{b}$. This is done with linear algebra techniques.

4. Form the action matrix M_a from the operation $a(\mathbf{x})\mathbf{b} = M_a^T\mathbf{b}$.

5. Calulate $\mathbf{b}$ from the eigenvalues of M_a^T and calculate the solutions from $\mathbf{b}$.

For this to be able to work, one generally needs to know how many solutions a system of polynomial equations has. If the matrix M_a^T does not have at least as many eigenvalues as solutions, we cannot hope to recover the different solutions. Thus, we need the candidate basis $\mathbf{b}$ to have a dimension that equals or is larger than the number of solutions. Selecting a too large candidate basis does in general not cause any problems besides a possible loss of performance, as all the recovered hypothetical solutions can be tried in (9) to remove extra false solutions.

Determination of the number of solutions to a polynomial system of equations have is done in this thesis with Macaulay2, [52], a software system for algebraic geometry. By conjecture, the number of solutions for the polynomial system is the same as for the polynomial system with random coefficients over the field $\mathbb{Z}_p$, where p is a large prime number, [53]. The number of solutions for this new system can be efficiently calculated with Macaulay using Gröbner basis techniques.

Several contributions have added to this method to buy numerical stability and computational efficiency to be able to solve a larger set of problems. In [54] and later in [55], steps 2 and 3 above are further studied to efficiently calculate the action matrix. In [56], symmetries in the initial polynomial systems of equations are exploited in steps 2 and 5.

5 Overview of the papers

This section gives a brief overview of the papers included in this thesis, as well as lists the author contributions.

Paper A

Simon Burgess, Yubin Kuang, and Kalle Åström, "TOA Sensor Network Self-Calibration for Receiver and Transmitter Spaces with Difference in Dimension", *Elsevier Signal Processing*, 2015.

In the first paper we solve the problem of finding both transmitter and receiver positions using only TOA measurements when there is a difference in dimensionality between the affine subspaces spanned by receivers and transmitters. Using linear techniques and requiring only minimal number of receivers and transmitters, an algorithm is constructed for general dimension p for the lower dimensional subspace. Degenerate cases are determined and partially characterized. The algorithm

is further extended to overdetermined cases. Utilizing the minimal solver, an algorithm using the RANSAC paradigm has been constructed to simultaneously solve the calibration problem and remove severe outliers. Simulated and real experiments show good performance for the minimal solver and the RANSAC-like algorithm under noisy measurements.

Author Contributions: KÅ suggested the topic of study. SB has worked and helped developed the initial idea, written most of the code, done the experiments, evaluations, developed the theoretical results, and writing for the journal version. YK and KÅ has helped with the initial idea, feedback on the direction of the work, and proof reading.

Paper B

Yubin Kuang, Erik Ask, Simon Burgess, and Kalle Åström, "Understanding TOA and TDOA Network Calibration using Far Field Approximation as Initial Estimate", *International Conference on Pattern Recognition Applications and Methods*, Algarve, Portugal, 2012.

In the second paper we present a study of the far field approximation to the problem of determining both the direction to a number of transmitters and the positions of the receivers, using TDOA or TOA measurements. In the far field approximation we assume that the distance between receivers are small in comparison to the distances to the transmitters from the receivers. The problem can be solved uniquely with at least four receivers and at least six real or virtual transmitters. The failure modes of the problem are studied and characterized. We also study to what extent the solution can be obtained in these degenerate configurations. The solution algorithm for the minimal case is extended to the overdetermined case in a straightforward manner. We also implement and test algorithms for nonlinear optimization of the residuals. In experiments we explore how sensitive the calibration is with respect to different degrees of far field approximations of the transmitters and with respect to noise in the data.

Author Contributions: KÅ and YK conceived the study. SB, EA, YK, and KÅ has jointly and in approximately equal shares written the code, designed and run the experiments, contributed to the theoretical contributions of the paper, and written the paper.

Paper C

Simon Burgess, Yubin Kuang, Johannes Wendeberg, Kalle Åström, and Christian Schindelhauer, "Minimal Solvers for Unsynchronized TDOA Sensor Network Calibration using Far Field Approximation", *International Symposium on Algorithms and Experiments for Sensor Systems, Wireless Networks and Distributed Robotics*, Sophia Antipolis, France, 2013.

In the third paper we extend the work of the second paper by presenting two novel approaches for the problem of self-calibration of network nodes using only TDOA when both receivers and transmitters are unsynchronized. We consider the previously unsolved minimum problem of far field localization in three dimensions, which is to locate four receivers by the signals of nine unknown transmitters, for which we assume that they originate from far away. The first approach uses that the time differences between four receivers characterize an ellipsoid. The second approach uses linear algebra techniques on the matrix of unsynchronized TDOA measurements. This approach is easily extended to more than four receivers and nine transmitters. Both simulated and a real experiment support the feasibility of the methods.

Author Contributions: SB and JW conceived and planned the study. SB and YK developed the matrix factorization method, and implemented it. JW and CS developed the ellipsoid method, and JW wrote the corresponding code. SB and JW planned and executed the experiments, developed the theoretical results, and wrote most of the paper with help from the other authors.

Paper D

Zhayida Simayijiang, Simon Burgess, Yubin Kuang, and Kalle Åström, "TOA-Based Self-Calibration of Dual-Microphone Array", *IEEE Journal on Selected Topics in Signal Processing*, 2015.

In the fourth paper we study the TOA based self-calibration problem of several dual microphone arrays for known and unknown rack distance, and also for affine space with different dimensions for receiver and sender spaces. Particularly we analyze the minimum cases and present minimum solvers for the case of microphones and speakers in 3D/3D, in 2D/3D, and in 3D/2D, with given or unknown rack length. We identify for each of these minimal problems the number

of solutions in general and develop efficient and numerically stable, non-iterative solvers. We demonstrate that the proposed solvers are numerically stable in synthetic experiments. We also demonstrate how the solvers can be used with the RANSAC paradigm. We apply our method for several real data experiments, using ultra-wide-band measurements and acoustic data.

Author Contributions: ZS and KÅ conceived and planned the study. The authors have all approximately equally implemented the methods. The paper was mainly written by ZS, but with extensive contributions from all other authors in approximately equal proportions. SB designed, executed, and evaluated the real life acoustic experiment.

Paper E

Yubin Kuang, Simon Burgess, Anna Torstensson, and Kalle Åström, "A Complete Characterization and Solution to the Microphone Position Self-Calibration Problem", *International Conference on Acoustics, Speech and Signal Processing*, Vancouver, Canada, 2013.

The fifth paper presents a solution to the problem of determining the positions of receivers and transmitters given all receiver-transmitter distances. We show for what cases such calibration problems are well-defined and derive closed-form, efficient, and numerically stable algorithms for the minimal TOA based self-calibration problems. Experiments on synthetic data show that the minimal solvers are numerically stable and perform well on noisy data. The solvers are also tested on two real datasets with good results.

Author Contributions: KÅ conceived and planned the study. YK and KÅ implemented and tested most of the methods, with help from SB and AT. The paper was mainly written by YK and KÅ, with contributions from the other authors. SB designed, executed, and evaluated the real life experiments.

Paper F

Simon Burgess, Mikael Högström, Björn Lindquist, and Kalle Åström, "Smartphone Positioning in Multi-Floor Environments Without Calibration or Added Infrastructure", to appear in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Madrid, Spain, 2016.

In the sixth paper we explore what can be done using existing WiFi-infrastructure and RSSI from these to smartphones, not using any calibration of the signal environment or manually set WiFi positions. We expand on previous work by using a multi-floor model taking into account dampening between floors, and optimize a target function consisting of least squares residuals, to find positions for WiFis and the smartphone measurement locations simultaneously. Pressure sensors are used to do floor estimation. The method is tested inside two multi-story buildings, with 5 stories each, with promising results.

Author Contributions: SB, KÅ and BL conceived and planned the study. All authors helped in approximately equal shares with implementing and testing of the methods. The paper was mainly written by SB and MH, with contributions from the other authors.

6 Topics for future research

The papers in this thesis mostly deal with calibration of sensor networks for receiver and transmitters when the measurements are already made. From this, one could imagine several paths for future research.

- One interesting direction is to employ the methods for complete systems, starting from signal acquisition. Indeed, this path has already been started. In [57], a complete system is described, starting from sound acquisition, continuing with using GCC-PHAT to do TDOA estimation, and finishing with calibrating the array and locating sources that have easily identifiable TDOA-measurements. In [58], WiFi round-trip time measurements are used to produce TOA measurements, which are then used to do calibration-free indoor localization. A continuation of this work could be to merge the calibration-free localization with a real-time application, to let both the receivers and transmitters move continuously. Here, fast minimal algorithms could be of great use, focusing on the latest measurements in a time series.

- It could be of great use for the research community to develop a set of guidelines or algorithms to determine when it is suitable to use the RANSAC paradigm and not. For instance to determine suitability of RANSAC from distributions and ratio of outliers and inliers. It is well known that RANSAC works well to get initial parameter estimation and outlier classification

when the measurements' outlier distribution is clearly different from a concentrated inlier distribution. To develop qualitative and quantitative methods to easily answer the suitability of RANSAC in comparison with, for instance, l_1-optimization or truncated l_1-optimization could be of great use.

- For indoor calibration-free localization using smartphones, COMBAIN and the Centre for Mathematical Sciences in Lund are looking in to deploying large-scale systems, and in the future providing commercial uses. As the methods in this thesis are mainly focused on one modality for calibration-free localization, the need for sensor fusion and real-time localization is prevalent. Sensor fusion is an active and exciting field, cf. [59], and is needed when dealing with large data sets from users with different platforms and hardware.

- Developing systems that automatically do calibration and then utilizing pre-existing research that do localization with calibrated sensor rigs could help emphasizing the utility of calibration-free localization. Such systems could also help to make calibration-free localization more robust and easier to use.

References

[1] Jonas Meldercreutz, "Om längders mätning genom dåns tilhielps," *Vetenskapsakademiens Handlingar*, vol. 2, pp. 73–77, 1741.

[2] Frank Parker Stockbridge, "How far off is that german gun?," *Popular Science*, vol. December, pp. 39, 1918.

[3] F. S. Hickson and V. E. Newton, "Sound localization," *The Journal of Laryngology & Otology*, vol. 95, pp. 29–40, 1 1981.

[4] Scott Gleason and Demoz Gebre-Egziabher, *GNSS Applications and Methods*, Artech House, ISBN: 9781596933293, 2009.

[5] Robin Henniges, "Current approaches of wi-fi positioning," in *IEEE Conference Publications*, 2012, pp. 1–8.

[6] Neal Patwari, Alfred O Hero, Matt Perkins, Neiyer S Correal, and Robert J O'dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on signal processing*, vol. 51, no. 8, pp. 2137–2148, 2003.

[7] Neal Patwari, Joshua N Ash, Spyros Kyperountas, Alfred O Hero, Randolph L Moses, and Neiyer S Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal processing magazine*, vol. 22, no. 4, pp. 54–69, 2005.

[8] Joseph Huang, David Millman, Morgan Quigley, David Stavens, Sebastian Thrun, and Alok Aggarwal, "Efficient, generalized indoor wifi graphslam," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1038–1043.

[9] Zheng Yang, Chenshu Wu, and Yunhao Liu, "Locating in fingerprint space: wireless indoor localization with little human intervention," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 269–280.

[10] He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury, "No need to war-drive: Unsupervised indoor localization," in *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, New York, NY, USA, 2012, MobiSys '12, pp. 197–210, ACM.

[11] Chenshu Wu, Zheng Yang, Yunhao Liu, and Wei Xi, "Will: Wireless indoor localization without site survey," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 4, pp. 839–848, Apr. 2013.

[12] Rijurekha Seni Krishna Kant Chintalapudi, Venkat Padmanabhan, "Zee : Zero-effort crowdsourcing for indoor localization," in *Mobicom*, August 2012.

[13] Pasi Pertilä, *Acoustic source localization in a room environment and at moderate distances*, Ph.D. thesis, Tampere University of Technology, 2009.

[14] Johannes Traa, "Multichannel source separation and tracking with phase differences by random sample consensus," 2013.

[15] B. Champagne, S. Bedard, and A. Stephenne, "Performance of time-delay estimation in the presence of room reverberation," *IEEE Transactions on Speech and Audio Processing*, vol. 4, no. 2, pp. 148–152, Mar 1996.

[16] T. Gustafsson, B. D. Rao, and M. Trivedi, "Source localization in reverberant environments: modeling and statistical analysis," *IEEE Transactions on Speech and Audio Processing*, vol. 11, no. 6, pp. 791–803, Nov 2003.

[17] M. D. Gillette and H. F. Silverman, "A linear closed-form algorithm for source localization from time-differences of arrival," *IEEE Signal Processing Letters*, vol. 15, pp. 1–4, 2008.

[18] K. C. Ho and M. Sun, "Passive source localization using time differences of arrival and gain ratios of arrival," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 464–477, Feb 2008.

[19] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 4, pp. 320–327, Aug 1976.

[20] J. H. DiBiase, H. F. Silverman, and M. S. Brandstein, "Robust localization in reverberent rooms," in *Microphone Arrays: Techniques and Applications*, M. Brandstein and D. Ward, Eds., pp. 157–180. Springer-Verlag, New York, 2001.

[21] Hoang Do, Harvey F Silverman, and Ying Yu, "A real-time srp-phat source location implementation using stochastic region contraction (src) on a large-aperture microphone array," in *2007 IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP'07*. IEEE, 2007, vol. 1, pp. I–121.

[22] X. Alameda-Pineda and R. Horaud, "A geometric approach to sound source localization from time-delay estimates," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 6, pp. 1082–1095, June 2014.

[23] Ivan Dokmanić, Reza Parhizkar, Andreas Walther, Yue M Lu, and Martin Vetterli, "Acoustic echoes reveal room shape," *Proceedings of the National Academy of Sciences*, vol. 110, no. 30, pp. 12186–12191, 2013.

[24] Axel Plinge, Florian Jacob, Reinhold Haeb-Umbach, and Gernot A Fink, "Acoustic microphone geometry calibration: An overview and experimental evaluation of state-of-the-art algorithms," *IEEE Signal Processing Magazine*, vol. 33, no. 4, pp. 14–29, 2016.

[25] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[26] Y. Kuang and K. Åström, "Stratified sensor network self-calibration from tdoa measurements," in *Proc. of the European Signal Processing Conference*, 2013.

[27] S. T. Birchfield and A. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 5, pp. 1025–1034, 2005.

[28] D. Niculescu and B. Nath, "Ad hoc positioning system (aps)," in *Proc. of Global Telecommunications Conference*, 2001.

[29] E. Elnahrawy, Xl. Li, and R. Martin, "The limits of localization using signal strength," in *Proc. of Sensor and Ad Hoc Communications and Networks*, 2004.

[30] V. C. Raykar, I. V. Kozintsev, and R. Lienhart, "Position calibration of microphones and loudspeakers in distributed computing platforms," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 1, pp. 70–83, 2005.

[31] M. Crocco, A. Del Bue, and V. Murino, "A bilinear approach to the position self-calibration of multiple sensors," *Signal Processing, IEEE Transactions on*, vol. 60, no. 2, pp. 660–673, 2012.

[32] J.C. Chen, R.E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *Signal Processing, IEEE Transactions on*, vol. 50, no. 8, pp. 1843–1854, 2002.

[33] S. Thrun, "Affine structure from sound," in *Proc. of Conference on Neural Information Processing Systems*, 2005.

[34] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, APR 2005.

[35] R. Biswas and S. Thrun, "A passive approach to sensor network localization," in *Proc. of International Conference on Intelligent Robots and Systems*, 2004.

[36] J. Wendeberg, F. Hoflinger, C. Schindelhauer, and L. Reindl, "Calibration-free tdoa self-localisation," *Journal of Location Based Services*, vol. 7, no. 2, pp. 121–144, 2013.

[37] Leonard Asimow and Ben Roth, "The rigidity of graphs, ii," *Journal of Mathematical Analysis and Applications*, vol. 68, no. 1, pp. 171–190, 1979.

[38] T. Eren, OK Goldenberg, W. Whiteley, Y.R. Yang, A.S. Morse, BDO Anderson, and PN Belhumeur, "Rigidity, computation, and randomization in network localization," in *Proc. of Conference of the IEEE Communications Society*, 2004.

[39] E.D. Bolker and B. Roth, "When is a bipartite graph a rigid framework," *Pacific Journal of Mathematics*, vol. 90, no. 1, pp. 27–44, 1980.

[40] M. Chasles, "Question 296," *Nouv. Ann. Math.*, vol. 14(50), 1855.

[41] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, ISBN: 0521540518, second edition, 2004.

[42] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[43] Peihua Li and Xianzhe Ma, "Robust acoustic source localization with tdoa based ransac algorithm," in *International Conference on Intelligent Computing*. Springer, 2009, pp. 222–227.

[44] Gergely Vakulya and Gyula Simon, "Fast adaptive acoustic localization for sensor networks," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 5, pp. 1820–1829, 2011.

[45] Olof Enqvist, Erik Ask, Fredrik Kahl, and Kalle Åström, *Robust Fitting for Multiple View Geometry*, pp. 738–751, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[46] Olof Enqvist, *Robust Algorithms for Multiple View Geometry: Outliers and Optimality*, Ph.D. thesis, Lund University, 2011.

[47] B. J. Tordoff and D. W. Murray, "Guided-mlesac: Faster image transform estimation by using matching priors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1523–1535, 2005.

[48] David A. Cox, John Little, and Donal O'Shea, *Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, 3/e (Undergraduate Texts in Mathematics)*, Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2007.

[49] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, 2005.

[50] David Nistér, "An efficient solution to the five-point relative pose problem," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 6, pp. 756–777, June 2004.

[51] Martin Byröd, "Numerical methods for geometric vision: From minimal to large scale problems," 2010.

[52] Daniel R. Grayson and Michael E. Stillman, "Macaulay2, a software system for research in algebraic geometry," Available at `http://www.math.uiuc.edu/Macaulay2/`.

[53] Y. Kuang, *Polynomial Solvers for Geometric Problems - Applications in Computer Vision and Sensor Networks*, Ph.D. thesis, Lund University, 2014.

[54] O. Naroditsky and K. Daniilidis, "Optimizing polynomial solvers for minimal geometry problems," in *2011 International Conference on Computer Vision*, Nov 2011, pp. 975–982.

[55] Yubin Kuang and Kalle Åström, *Numerically Stable Optimization of Polynomial Solvers for Minimal Problems*, pp. 100–113, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012.

[56] Yubin Kuang, Yinqiang Zheng, and Karl Åström, "Partial symmetry in polynomial systems and its application in computer vision," 2014.

[57] Zhayida Simayijiang, Fredrik Andersson, Yubin Kuang, and Kalle Åström, "An automatic system for microphone self-localization using ambient sound," in *European Signal Processing Conference (Eusipco 2014)*. EURASIP (European Association for Signal Processing), 2014, p. 5.

[58] Kenneth Batstone, Magnus Oskarsson, and Kalle Åström, "Robust time-of-arrival self calibration and indoor localization using wi-fi round-trip time measurements," in *2016 IEEE International Conference on Communications Workshops (ICC)*. IEEE, 2016, pp. 26–31.

[59] Fredrik Gustafsson, *Statistical sensor fusion*, Studentlitteratur,, 2010.

A

Paper A

TOA Sensor Network Self-Calibration for Receiver and Transmitter Spaces with Difference in Dimension

Simon Burgess, Yubin Kuang, Kalle Åström

Centre for Mathematical Sciences, Lund University, Lund, Sweden

Abstract

We study and solve the previously unstudied problem of finding both transmitter and receiver positions using only time of arrival (TOA) measurements when there is a difference in dimensionality between the affine subspaces spanned by receivers and transmitters. Anchor-free TOA network calibration has uses both in radio, radio strength and sound applications, such as calibrating ad hoc microphone arrays. Using linear techniques and requiring only minimal number of receivers and transmitters, an algorithm is constructed for general dimension p for the lower dimensional subspace. Degenerate cases are determined and partially characterized as when receivers or transmitters inhabits a lower dimensional affine subspace than was given as input. The algorithm is further extended to overdetermined cases in a straightforward manner. Utilizing the minimal solver, an algorithm using the Random Sample Consensus (RANSAC) paradigm has been constructed to simultaneously solve the calibration problem and remove severe outliers, a common problem in TOA applications. Simulated experiments show good performance for the minimal solver and the RANSAC-like algorithm under noisy measurements. Two indoor environment experiments using microphones and speakers gives a RMSE of 2.35 cm and 3.95 cm on receiver and transmitter positions compared to computer vision reconstructions.

Key words: TOA, array calibration, minimal problem, ad hoc microphone arrays.

1 Introduction

Sound ranging or sound localization has been used since World War I to determine the sound source using a number of microphones at known locations and measuring the time difference of arrival of sounds. The same mathematical model is today used both for applications based on acoustics, radio, signal strength or time-based information such as time of arrival (TOA) or time differences of arrival (TDOA), or a combination thereof. Although such problems have been studied extensively in the literature in the form of localization of e.g. a sound source using a calibrated receiver array, see e.g. [1–3], or localizing a source and refining an already good initial guess of a receiver array, see e.g. [4–6], the problem of calibration of a sensor array with unknown node positions using only measurements has received comparatively less attention.

In this paper we study and solve the previously unstudied sensor network calibration problem using only TOA measurements where there is a difference in dimensionality between the affine subspaces spanned by transmitters and receivers. This is of both theoretical and practical importance. We provide theorems to show when the problem, related to rigidity of bipartite graphs, is solvable. Applications can be in indoor navigation where receivers often are moving in a plane, e.g. a floor, whereas transmitters can be arbitrarily located in 3D, or for calibrating ad hoc microphone or antenna arrays. The case of array calibration when TOA measurements are available is also of importance in the corresponding TDOA calibration, where a stratified approach of first determining offsets and then solving the TOA calibration problem is a common approach, see e.g. [7, 8].

We prove that the TOA calibration problem for difference in dimension can be solved in closed form using linear techniques and give a solution scheme for general dimensionality p. Furthermore, the solver is extended to overdetermined cases. Simulated and real experiments support the feasibility of the method. The problem relates to the study of sensor networks under rigid graph theory [9, 10] where general graph structure is of interest. The TOA self-calibration problem studied here corresponds to a special case - bipartite graphs [11].

Several previous contributions dealing with sensor network calibration rely on prior knowledge or extra assumptions about locations of the sensors to initialize the problem, see e.g. [12–17]. Here we propose that being able to solve the self-calibration problem without using more than the TOA measurements between transmitters and receivers opens up for interesting applications, e.g. calibration of radio and acoustic sensor networks on the fly, determining reflections of receivers

and transmitters while moving in an unknown terrain.

In [18] and refined in [19] a far field approximation was utilized to solve the TOA and TDOA case. Initialization of TOA calibration using only measurements has been studied in [20, 21], where solutions to the minimal cases of three transmitters and three receivers in the plane, or six transmitters and four receivers in 3D are given. Here, minimal case means using the fewest possible number of receivers and transmitters for the problem to have finite positive number of solutions. Calibration of TDOA networks was studied in [7] and further improved upon in [8]. In [22, 23], a TDOA setup is used for indoor navigation based on non-linear optimization, but the methods can get stuck in local minima and are dependent on initialization.

Of the above contributions, [7, 8, 18–21] can be said to solve a calibration problem using only measurements with either minimal or close to minimal data. Studying minimal cases for sensor network calibration are both of theoretical importance and essential to develop fast stable algorithms suitable in random sample consensus (RANSAC) [24] and other schemes. RANSAC schemes can be used to weed out outliers in erroneously matched data which is a common problem in TOA and TDOA applications. The difference in dimensionality problem we study here is either a degenerate case for the papers above, or requires estimating several extra variables because of the assumption that receivers and transmitters lie in a same dimensional subspace. Thus the previous methods are ill suited or cannot be applied to the problem we study here.

1.1 Problem Formulation

In this paper we study the so called TOA node self-calibration problem when the dimension of the affine subspaces spanned by receivers and transmitters are different. We assume that (i) the speed of the medium v is known, and thus all time measurements are transformed to distances by multiplication by v, (ii) receivers can distinguish which TOA signal comes from which transmitter. This can be done in practice by e.g. separating the signals temporally or by any other signal characteristics such as frequency. For instance, if one has several speakers emanating sounds at the same time but at different frequency bands, separation can readily be done with Fourier transform, and TOA measurments can be calculated by the difference of time of sounds emanating and sound onset in microphones. In radio received signal strength indication applications, which gives distance measurements from base stations to a receiver, the separation is done by

coding techniques.

Problem 1.1. Assume receiver coordinates $\mathbf{r}_i, i = 1, \ldots, m$ are embedded in a p-dimensional affine space Π and transmitter coordinates $\mathbf{s}_j, j = 1, \ldots, n$ are embedded in a d_2-dimensional, possibly infinite, space Π_2, where $\Pi \subset \Pi_2$. Given absolute distance measurements $d_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2$, determine $\mathbf{r}_i$ and $\mathbf{s}_j$.

Without loss of generality, receivers and transmitters can be interchanged. The problem is somewhat related to the classical multidimensional scaling (MDS) problem, with two major differences: i) We here only have distance measurements between receivers and transmitters, whereas MDS has distances between all points. ii) The underlying dimensions of the receiver and transmitter spaces differ. We need to consider what we can expect of a solution. First of all, while still fulfilling the distance measurements, a solution can be changed according to

$$\begin{aligned} \bar{\mathbf{r}}_i &= \mathbf{r}_i + \mathbf{t}, \ \bar{\mathbf{s}}_j = \mathbf{s}_j + \mathbf{t}, \\ \bar{\mathbf{r}}_i &= R\mathbf{r}_i, \ \bar{\mathbf{s}}_j = R\mathbf{s}_j, \end{aligned}$$

where R is a rotation and/or mirroring matrix and $\mathbf{t}$ is a translation. Assuming a rotation so that the last coordinate of the receivers $\mathbf{r}_i$ are 0, the problem has undeterminable degrees of freedom (gauge freedom) of p from the translation and $p(p-1)/2$ from the rotation, which gives a total number of $p + p(p-1)/2$.

Furthermore, for the transmitters $\mathbf{s}_j$ one can only hope to determine the orthogonal projection onto Π and the distance to Π, as a rotation around Π will keep the receivers constant but change the transmitters, without changing the measurements. Therefore, for practical purposes one can assume that transmitters are embedded in a $p+1$ dimensional space. Assuming this is the case, each transmitter can still be mirrored in Π without changing the distances.

We denote the problem as minimal if the number of solutions $\mathbf{r}_i, \mathbf{s}_j$ for generic distance measurements d_{ij} is finite, and at least one on a non-zero measure subset of the measurement space $\mathbb{R}^{m \times n}$. A solution is considered to be the same as another solution if and only if it only differs up to translation, rotation, mirroring in coordinate axis and mirroring of each $\mathbf{s}_j$ in Π.

There are many ways to account for these ambiguities. We choose to translate and rotate the coordinate system so that: (i) The translation is locked by setting $\mathbf{r}_1 = \mathbf{0}$. (ii) The mirroring is locked by setting $\mathbf{s_j}$'s last coordinate to positive. (iii) The last coordinate of $\mathbf{r}_i$ is 0. The extra coordinate for transmitter j will thus be

$\mathbf{s_j}$'s last coordinate. The rest of the rotation ambiguity is left free, and fixed first when comparing two solutions using [25].

2 Theory

Here follows a detailed description to the solver for the TOA node self-calibration problem. We first develop a minimal algorithm, where the noiseless case is naturally considered, as we have just enough measurements to solve the problem. Thus there is no way to compensate for noisy measurements without assuming further priors such as known bias. Let $D = \left[d_{ij}\right]_{m\times n}$ be the matrix with the distance measurements, and $D_2 = \left[d_{ij}^2\right]_{m\times n}$ be the matrix with the distance measurements squared. As $d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j) = \mathbf{r}_i^T\mathbf{r}_i - 2\mathbf{r}_i^T\mathbf{s}_j + \mathbf{s}_j^T\mathbf{s}_j$, the squared distances can be written as

$$d_{ij}^2 = \begin{bmatrix}1 & -2\mathbf{r}_i^T & \mathbf{r}_i^T\mathbf{r}_i\end{bmatrix}\begin{bmatrix}\mathbf{s}_j^T\mathbf{s}_j & \mathbf{s}_j^T & 1\end{bmatrix}^T = E_iF_j.$$

This leads to that $D_2 = EF$ where E is a matrix with E_i as the i^{th} row and F is a matrix with F_j as the j^{th} column. This factorization was used in [7]. As the last coordinate of $\mathbf{r}_i$ only has zeros, the second last column of E will be zeros, multiplying with the second last row of F. We can thus remove this row and column from the expression, forming

$$\begin{aligned} d_{ij}^2 = \begin{bmatrix}1 & -2\bar{\mathbf{r}}_i^T & \bar{\mathbf{r}}_i^T\bar{\mathbf{r}}_i\end{bmatrix}\begin{bmatrix}\mathbf{s}_j^T\mathbf{s}_j & \bar{\mathbf{s}}_j^T & 1\end{bmatrix}^T = \bar{E}_i\bar{F}_j \Rightarrow \\ D_2 = \bar{E}\bar{F}, \end{aligned} \tag{1}$$

where $\bar{\mathbf{r}}_i, \bar{\mathbf{s}}_j$ are $\mathbf{r}_i, \mathbf{s}_j$ with the last coordinate removed, $\bar{E}$ is an $m \times \left(p+2\right)$ matrix and $\bar{F}$ is an $(p+2)\times n$ matrix. This tells us that the rank of D_2 is at most $p+2$.

We form the matrix $S = \left[s_{ij}\right]_{(m-1)\times n}$, $i = 1, \ldots, m-1$

$$\begin{aligned} s_{ij} &= d_{i+1,j}^2 - d_{1,j}^2 \\ &= (\mathbf{r}_{i+1} - \mathbf{s}_j)^T(\mathbf{r}_{i+1} - \mathbf{s}_j) - (\mathbf{r}_1 - \mathbf{s}_j)^T(\mathbf{r}_1 - \mathbf{s}_j) \\ &= \bar{\mathbf{r}}_{i+1}^T\bar{\mathbf{r}}_{i+1} - 2\bar{\mathbf{r}}_{i+1}^T\bar{\mathbf{s}}_j \\ &= \begin{bmatrix}-2\bar{\mathbf{r}}_{i+1}^T & \bar{\mathbf{r}}_{i+1}^T\bar{\mathbf{r}}_{i+1}\end{bmatrix}\begin{bmatrix}\bar{\mathbf{s}}_j^T & 1\end{bmatrix}^T = \tilde{E}_i\tilde{F}_j, \end{aligned} \tag{2}$$

as $\mathbf{r}_1 = \mathbf{0}$ and $\mathbf{r}_i$'s last coordinate is 0. By using the first receiver we have eliminated the quadratic constraints on the transmitters, effectively forming equations in our unknowns $\mathbf{r}_i$ and $\mathbf{s}_j$ which only depends on the p first coordinates, i.e. their orthogonal projection in Π. The equations are also linear in $\mathbf{s}_j$.

We note that $S = \tilde{E}\tilde{F}$, where $\tilde{E}$ is an $(m-1) \times (p+1)$ matrix where the i^{th} row is $\tilde{E}_{i+1}$ and $\tilde{F}$ is an $(p+1) \times n$ matrix where the j^{th} column is $\tilde{F}_j$. This tells us that S is at most of rank $p+1$.

2.1 Solving in Π

We seek a factorization $S = \tilde{E}\tilde{F}$ such that $\tilde{F}$ has a last row of ones, and the quadratic constraints in each row of $\tilde{E}$ is fulfilled (2). Assuming that S has rank $p+1$, we start by doing the compact singular value decomposition (svd)

$$S = \underbrace{U}_{X_0} \underbrace{\Sigma V^T}_{Y_0} = X_0 Y_0, \tag{3}$$

where U, Σ and V are $(m-1) \times (p+1)$, $(p+1) \times (p+1)$ and $n \times (p+1)$ respectively. We continue by expanding

$$S = \underbrace{X_0 B_0^{-1}}_{X_1} \underbrace{B_0 Y_0}_{Y_1} = X_1 Y_1, \tag{4}$$

where B_0's last row is chosen so that Y_1 will have a last row of ones. This is done by solving a linear system of equations with $p+1$ unknowns, which tells us that we need $p+1$ transmitters so that we can solve the system uniquely. This can always be done as Y_0 has full rank, implied by the assumption that S has rank $p+1$.

The other p rows of B_0 are calculated with Gram-Schmidt orthogonalization so that B_0 is a unitary real matrix times a scalar. To do this, take the top p rows to be p different rows of the identity matrix of dimension $p+1$ so that all rows are linearly independent, and then apply the Gram-Schmidt process starting from the last row. Then the condition number of B_0 will be 1 so that $X_0 B_0^{-1}$ does not lose unnecessary precision. Y_1 does now have the right properties suggested by

(2) but X_1 does not. We continue by expanding

$$S = \underbrace{X_1 B_1}_{\tilde{E}} \underbrace{B_1^{-1} Y_1}_{\tilde{F}} = \tilde{E}\tilde{F},\ B_1 = \begin{bmatrix} A & \mathbf{b} \\ \vec{0} & 1 \end{bmatrix}, \tag{5}$$

where A is a general invertible $p \times p$ matrix, $\mathbf{b}$ is a general $p \times 1$ vector and $\vec{0}$ is a vector of zeros. This implies that B_1^{-1} has the same restrictions as B_1 which gives the most general form for preserving the last row of ones and rank in Y_1 when forming $B_1^{-1} Y_1$. It remains to determine A and $\mathbf{b}$.

This is done by enforcing the quadratic constraints of the rows of the left matrix in (2) on $X_1 B_1$. We denote row i of X_1 as $\begin{bmatrix} \mathbf{v}_i & v_{i,p+1} \end{bmatrix}$ where $\mathbf{v}_i$ is a vector of length p and $v_{i,p+1}$ is the last element. The constraints then become

$$\mathbf{v}_i A A^T \mathbf{v}_i^T = 4\left(\mathbf{v}_i \mathbf{b} + v_{i,p+1}\right). \tag{6}$$

As $AA^T = ARR^T A$ for any rotation and/or mirroring matrix R, and this R equates exactly to a rotation/mirroring of $\mathbf{r}_i$ and $\mathbf{s}_j$ in Π, we need only to solve (6) for the symmetric matrix $C = A^T A$. This gives $p + p(p+1)/2$ unknowns from $\mathbf{b}$ and C, which tells us that we need $1 + p + p(p+1)/2$ receivers to be able to solve the linear equations (6) in the unknowns C and $\mathbf{b}$ uniquely, as well as the system to have full rank. The extra receiver comes from losing a row in S by subtracting the fist row of D_2 in (2).

When C and $\mathbf{b}$ have been determined, an A can be calculated from e.g. cholesky factorization in C. Now the left and right hand factorization of S, $\tilde{E} = X_1 B_1$ and $\tilde{F} = B_1^{-1} Y_1$ can be calculated. $\tilde{\mathbf{r}}_{i+1}$ can then be calculated as the first p elements of row i of $\tilde{E}$ divided by -2, and $\tilde{\mathbf{s}}_j$ are the first p elements in respective column of $\tilde{F}$, according to (2). The last coordinate $\mathbf{s}_j$ can then be recovered by using (1) and the last coordinate of $\mathbf{r}_i$ is 0. A summary of the steps can be seen in Algoritm 1.

2.2 Minimal cases

It can be seen from the following theorem that the algorithm solves a minimal problem, defined in Section 1.1.

Theorem 2.1. *When having $m = 1 + p + p(p+1)/2$ receivers and $n = p + 1$ transmitters, Problem 1.1 is minimal. Furthermore there is only one solution up*

Algorithm 1 Minimal solver for the TOA calibration problem

Input: Dimension p of receiver space, TOA measurement matrix $D = \left[d_{ij}\right]$ of size $m = 1 + p + p(p+1)/2$ by $n = p + 1$.
Output: Receiver and transmitter positions $\mathbf{r}_i$ and $\mathbf{s}_j$.
Postconditions: (i) S has rank $p+1$, (ii) linear system $\mathbf{v}_i C \mathbf{v}_i^T = 4\left(\mathbf{v}_i \mathbf{b} + v_{i,p+1}\right)$ has full rank, (iii) C is positive definite.

1: Set $S := \left[d_{i+1,j}^2 - d_{1,j}^2\right]$
2: Calculate the svd $S = U\Sigma V^T$ and set X_0 to first $p+1$ columns of U and Y_0 to first $p+1$ rows of ΣV^T
3: Calculate B_0 such that $B_0 Y_0$ has last row of 1's and such that B_0 is unitary times a scalar
4: Set $X_1 := X_0 B_0^{-1}$, $Y_1 := B_0 Y_0$, $\mathbf{v}_i$ to the first p elements of row i of X_1 and $v_{i,p+1}$ to the last.
5: Solve for the unknowns in the symmetric matrix C and vector $\mathbf{b}$ the $m-1$ linear equations $\mathbf{v}_i C \mathbf{v}_i^T = 4\left(\mathbf{v}_i \mathbf{b} + v_{i,p+1}\right)$
6: Calculate the cholesky decomposition $C = AA^T$ and set $B_1 := \begin{bmatrix} A & \mathbf{b} \\ \vec{0} & 1 \end{bmatrix}$
7: Set $\tilde{E} := X_1 B_1$, $\tilde{F} := B_1^{-1} Y_1$, $\tilde{\mathbf{r}}_{i+1}$ to the p first elements in row i of $\tilde{E}$ divided by -2, and $\tilde{\mathbf{s}}_j$ to the first p elements of column j of $\tilde{F}$
8: Solve for $s_{j,z}$ to the positive sign solution, $d_{1j}^2 = \left[\tilde{\mathbf{s}}_j \; s_{j,z}\right]^T \left[\tilde{\mathbf{s}}_j \; s_{j,z}\right]$
9: Set $\mathbf{r}_1 := \vec{0}$, $\mathbf{r}_i := \left[\tilde{\mathbf{r}}_i \; 0\right]$ and $\mathbf{s}_j = \left[\tilde{\mathbf{s}}_j \; s_{j,z}\right]$

to translation and rotation/mirroring in Π, and mirroring of the last transmitter coordinate in Π.

Proof. Looking at the steps of Algorithm 1, we see that we do not lose any solutions due to specific choices of parameters, except equivalent solutions up to gauge freedom and mirroring in Π, according to Section 1.1. Thus, we have only one solution, and the problem is minimal. □

A common, but sometimes misleading, way of make likely that the problem is minimal is by counting the degrees of freedom of the problem:

$$\mathrm{DOF}(m,n) = pm + (p+1)n - p - p(p-1)/2. \tag{7}$$

The positive terms come from the coordinates of $\mathbf{r}_i$ and $\mathbf{s}_j$ respectively, and the negative terms come from the gauge freedom in Section 1.1. We have mn measurements, whereas the algorithm needs that $m = 1+p+p(p+1)/2$ and $n = p+1$, giving $mn = (1+p+p(p+1)/2)(p+1)$ measurements. Comparing the number of measurements to the degrees of freedom, we see that they are equal. However, this by itself is not a sufficient condition for minimality. Note that there are other constellations of m, n where the mn matches the degrees of freedom in (7). For instance, if $p = 2$, we have $m = 4, n = 5$, and for $p = 3$ we have three other constellations. In general, while they may at first glance look like good candidate constellations for a minimal case to solve the calibration problem, this is not the case. When both $m, n > p+1$, the matrix compaction (2) tells us that the matrix S does not have full rank, as $Rank(S) = p+1$. Thus, some of the transformed measurements in S can be derived from the others and are not left free. There are in some sense fewer measurements than mn due to the constraints imposed by the problem for these constellations.

Assuming that the receivers and transmitters are in a same dimensional subspace with dimension $p+1$, we can use the algorithm to just solve in Π, skipping step 8 and 9 in the algorithm. Comparing degrees of freedom to measurements now suggests an overdetermined system. Solving this problem in 3D we need 10 receivers and 4 transmitters, which is on par with [7] that also utilizes linear techniques. To solve the minimal problem in 3D, e.g. using 6 receivers and 4 transmitters, requires solving systems of polynomial equations [21].

2.3 Degenerate cases

There are special cases of constellations for receivers and transmitters where the algorithm cannot produce a solution, for example if two transmitters are placed on top of each other. Then the measurements of the corresponding transmitters become duplicates, and we effectively have one transmitter less than needed. These special cases, denoted degenerate cases, are characterized by the following theorem.

Theorem 2.2. *Degenerate cases for the minimal algorithm are when i) the assumption that S is of full rank, i.e. $p+1$, does not hold, ii) step 5 gives a linear system that does not have full rank, or iii) C is not of full rank.*

i) happens if and only if X_0 or Y_0 in the svd in step 2 is rank deficient. Y_0 is rank deficient if and only if the projection onto Π of the transmitters, $\tilde{\mathbf{s}}_j$, lie in an even

lower dimensional affine subspace than Π. X_0 is rank deficient if receivers $\mathbf{r}_i$ span a lower dimensional subspace than Π.

Proof. Looking at the assumptions for the algorithm to work, the only ones are the ones mentioned in (i), (ii) and that C should be positive definite. When C is of full rank, but not positive definite, the special case after the proof characterizes this. When C is not of full rank, there is no full rank matrix A such that $C = A^T A$, as in that case C would have full rank. Thus i), ii) and iii) constitute the degenerate cases.

The matrix S is trivially rank deficient if and only if X_0 or Y_0 is. Y_0 is rank deficient if and only if $\tilde{F}$ is rank deficient, as $\tilde{F}$ is Y_0 multiplied with invertible matrices. $\tilde{F}$ is rank deficient if and only if $\exists\ a_1, \ldots, a_{p+1} \in \mathbb{R}$, not all $= 0$, such that

$$\sum_{i=1}^{p} a_i \tilde{F}_i = a_{p+1} \vec{1} \tag{8}$$

as $\tilde{F}$ has a last row of ones. As each row i of $\tilde{F}$, $\tilde{F}_i$, consists of the i^{th} coordinates of $\tilde{\mathbf{s}}_j$, (8) is equivalent to that $\tilde{\mathbf{s}}_j$ lie in an affine space with dimension $< p$.

The matrix X_0 is rank deficient if and only if $\tilde{E}$ is rank deficient, as $\tilde{E}$ is X_0 multiplied with invertible matrices. $\tilde{E}$ is rank deficient if the first p columns do not span a p-dimensional space. As column k contains the k^{th} coordinates for $\mathbf{r}_i$, this gives that $\mathbf{r}_i$ spans a lower dimensional subspace than Π. □

Note that the degenerate cases characterized in i) and iii) are inherent to the problem, not the algorithm. If the receivers or transmitters lie in a lower dimensional subspace than assumed, there are fewer degrees of freedoms to estimate than assumed. If the matrix C is rank deficient, there exist no full rank linear transformation A such that the constraints in (2) are fulfilled.

A special case, resulting in complex solutions, is when the resulting matrix C in step 5 is not positive definite, but full rank. If this happens, then there exists no real solution for A such that $C = AA^T$. A will have to be complex, and can be solved by eigenvalue decomposition $C = Q^T D Q = Q^T \sqrt{D}\sqrt{D} Q = AA^T$, giving complex solutions $\mathbf{r}_i$ and $\mathbf{s_j}$ fulfilling the measurements in D. This can also happen for the $p+1^{\text{th}}$ coordinate in $\mathbf{s}_j$, if the projection $\tilde{\mathbf{s}}_j$ is larger than d_{1j}^2, which is used to calculate the last coordinate in step 8.

2.4 More receivers and transmitters

When having more than $1 + p + p(p+1)/2$ receivers, $p+1$ transmitters and measurements d_{ij} possibly corrupted by noise, the algorithm can be expanded by i) approximating S in (3) by the closest rank $p+1$-matrix in Frobenius norm, by setting the singular values after the $p+1$ first to zero, ii) taking the linear least squares solution to the system of equations resulting from (4) and (6) respectively. This will not be an optimal solution in any formal sense, but will give a good solution which can serve e.g. as an initial estimate for further non-linear optimization techniques.

The computational complexity has an upper bound of $O(n^2m + nm^2 + n^3 + m^3)$, where the first three terms come from the svd in step 2, and the last one from solving the $m-1$ linear equations in step 5.

From here on the extended algorithm will be used. Note that when having only the minimum numbers of receivers and transmitters, the extended and minimal solver are equivalent.

2.5 A RANSAC-like algorithm

Outliers in the measurements d_{ij} is a common problem in applications, usually occuring from erroneous matching of event detections from e.g. microphones or antennas. To handle this, we have developed a RANSAC-like scheme. RANSAC was first used in image analysis [24]. The advantages of the proposed algorithm is that it can estimate receiver and transmitter positions and at the same time weed out outliers in the measurement set. In practice, RANSAC-schemes has been used to great success in image analysis and computer vision, see [26] for several examples.

At the core, the RANSAC-like algorithm presented here uses the minimal solver presented in Section 2. Here, the minimality is crucial. The RANSAC-like scheme needs to get a subset of all inliers for at least one of its iterations. The fewer measurements it needs to fit an initial model, the higher the chance of that model is fit with inliers.

Here follows the general outline of the algorithm [1]. The algorithm given here is for clarity the case when receivers are in 2D and transmitters in 3D, but it is easily generalized. Given a set of measurements $\{d_{ij}\}$, we want to find receiver

[1] The algorithm and all code is available at
`https://github.com/SimonEBurgess/TOA-self-calibration`

and transmitter positions and at the same time weed out outliers. Pick a minimal subset of measurements, in this case corresponding to 6 receivers and 3 transmitters. Fit a model using the minimal algorithm in Section 2. Now we have positions for 6 receivers and 3 transmitters. From these 6 receivers, trilaterate all the transmitter positions. Now we count the number of total inliers in the set of measurements we have receiver and transmitter positions to match, defining d_{ij} as an inlier if

$$|d_{ij} - ||\mathbf{r}_i - \mathbf{s}_j||_2| \leq \varepsilon \tag{9}$$

for a given ε. If the initial minimal set of measurements were all inliers, and the trilaterations for all transmitters were done robustly, this would give a good estimation of 6 receiver, all transmitter positions and which of the measurements we have receiver and transmitter positions for that are inliers. From here, we can trilaterate the missing receivers, and if this is done robustly, it is now easy to use (9) to classify any measurement as inlier or outlier.

Summarizing the assumptions made we have two problems: i) How do we know the initial set of measurements are inliers? ii) How do we robustly trilaterate more receivers and transmitters in the presence of outliers? In RANSAC approaches, the key argument for problem i) is that when having a large amount of data and repeating the whole process many times, the probability of randomly selecting an initial subset that are all inliers at least once, approches one for a given percentage of outliers. Thus, we iterate the process of solving the minimal problem, trilaterating all transmitters and counting inliers. Each time we randomly select the initial minimal set of measurements and then trilaterate up all transmitters. Over all runs, we save the final result that has the most number of inliers. Each of these iterations is called a main iteration. As we need to iterate the whole process, this explains the need to have a fast minimal solver at the core. Problem ii) is dealt with in a similar manner. For instance, when trilaterating a new receiver from known transmitters, we do it several times using randomly selected minimal number of measurements each time. Then we use the receiver position $\mathbf{r}_k$ that gives the most number of inliers on the set $\{d_{kj} : j \in \text{known transmitter index}\}$.

As transmitters are trilaterated using only 6 receiver positions, even though we try to find the transmitter position that fits the most number of the 6 measurements, this might not work well if we have many outliers in these six measurements. Fortunately we can detect this by seeing that for these indices $\{l\}$, we should have very few inliers in $\{d_{il} : i = 1, 2, \ldots, m\}$ when having completed the process of estimating all receiver and transmitter positions. Thus, after

Algorithm 2 RANSAC-scheme for finding node positions and inliers

Input: Dimension p, TOA measurement matrix $D = \left[d_{ij}\right]$, *nbrIterMain*, *nbrIterTril*, Tolerance for inlier classification ε, Max # outliers for not retrilaterating transmitter *nbrOutliersTol*.

1: **for** $i = 1$ **to** *nbrIterMain* **do**
2: Randomize $1 + p + p(p+1)/2$, $p+1$ different receiver and transmitter indices. Run Algorithm 1 using the corresponding part of D.
3: If solution is real, use the obtained receivers to trilaterate the unknown transmitters. For each new transmitter, iterate over every combination of $p+1$ receivers from the $1 + p + p(p+1)/2$ known. For each iteration, use the $p+1$ receivers to trilaterate the transmitter. Save the transmitter position that gave the most inliers for the $1+p+p(p+1)/2$ corresponding measurements, according to (9).
4: Check number of inliers, using (9), for the part of D with known receivers and transmitters. If the best so far, save receiver and transmitter positions.
5: **end for**
6: For each unknown receiver, trilaterate the position using random $p+1$ transmitters. Do this *nbrIterTril* times. Save the receiver position that has most inliers according to (9).
7: For each transmitter that has more outliers than *nbrOutliersTol* on the corresponding measurements, retrilaterate it using random $p+1$ receivers. Retrilaterate the transmitter position *nbrIterTril* times. Save the transmitter positions that has most inliers according to (9).
8: Check what measurements are inliers using (9).

all receiver and transmitter positions have been trilaterated, we re-trilaterate the transmitters that correspond to few inliers using our previous scheme, now having all the receivers.

The same problem is not present when trilaterating the receiver positions, as we then have all the transmitter positions to trilaterate from. Some of the transmitters might be wrong at this step, but as we search for each receiver iteratively using only a minimal number of transmitters each time, this does not generally cause any problems.

3 Results

3.1 Simulated Experiments

For all experiments in this section, ground truth receivers $\mathbf{r}_{i,gt}$ and transmitters $\mathbf{s}_{j,gt}$ are simulated, and from there a distance matrix $D_{gt} = [d_{ij}]$ has been calculated. D_{gt}, sometimes corrupted with noise, missing data and outliers, serves as input.

The receivers, residing in the subspace Π of dimension p, has been drawn from a uniform distribution over a unit cube centered around the origin. The transmitters $\mathbf{s}_{j,gt}$, residing in Π_2 with dimension $p+1$, have their distances from the origin drawn from a uniform distribution $U(0.1\sqrt{p}, \sqrt{p})$ and then uniformly distributed over the sphere with that radius.

To be able to evaluate the quality of the solution, $\mathbf{r}_i$, $\mathbf{s}_j$ are rotated, mirrored and translated so that $\sum_i ||R(\mathbf{r}_i - \mathbf{t}) - \mathbf{r}_{i,gt}||_2^2 + \sum_j ||R(\mathbf{s}_j - \mathbf{t}) - \mathbf{s}_{j,gt}||_2^2$ is minimized, where R and t is a rotation/mirroing and translation in Π respectively. Finding R and $\mathbf{t}$ is done by using [25]. The relative error of the solution is then calculated as $||\mathbf{r}_i - \mathbf{r}_{i,gt} \;\; \mathbf{s}_j - \mathbf{s}_{j,gt}||_{Fro} / ||\mathbf{r}_{i,gt} \;\; \mathbf{s}_{j,gt}||_{Fro}$ where $||\cdot||_{Fro}$ denotes the Frobenius norm. All experiments were run on a desktop computer using Intel Core 2 Duo CPU with two 2.8 GHz processors, implemented in Matlab.

For the minimal solver, 1000 experiments were run for $p = 2, 3, 4$ each. Histograms of relative errors can be seen in Figure 1. The mean computational time of the algorithm over these runs was 3.0 ms.

In Figure 2 the mean relative errors of 1000 runs, $p = 2$, for different number of receivers and transmitters, are plotted against the standard deviation of additive white Gaussian noise on the measurements D_{gt}. The errors are only calculated for the real solutions, as the complex ones do not have a physical meaning, though they fulfill the measurements D if not overdetermined, and approximately fulfill them if overdetermined. Complex solutions correspond to the special cases described in Section 2.3.

To see the validity of using the minimal solver to weed out outliers and simultaneously find an effective solution to the calibration problem, the RANSAC-like algorithm in Section 2.5 was run on measurements D_{gt}. For each experiment, each measurement has a independent chance of being replaced by outliers or missing data. A measurement d_{ij} that was replaced by an outlier, was redrawn from a uniform distribution $U(0, \sqrt{p})$. As it is generally harder to do an effective reconstruction from partly erroneous data than partly missing data, the chance of a measurement being removed as missing data is kept to a moderate 1% over all ex-

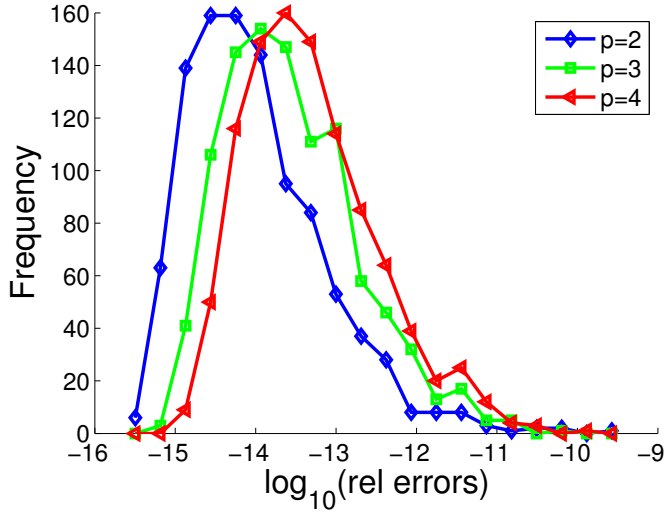


Figure 1: Histograms of relative errors for the minimal solver. Here p is the dimension of the affine space the receivers span.

periments. The dimension of the subspace receivers is set to $p = 2$. In Figure 3a, histograms relative reconstruction can be seen for different percentage of missing data, for 100 runs each. 10 receivers and 15 transmitters were used, and parameters are set to 300 main iterations, 50 iterations for trilaterating each receiver, inlier tolerance classifier ε of 0.005 and a tolerance of 3 outliers per column to not retrilaterate transmitters at the end. In Table 1, the average relative frequency of erroneously classified measurements as outliers or inliers can be seen, together with outlier false negative and false positive. Average execution time was 1.1s.

The same experiment was made but with additive Gaussian noise added to the measurements, with a standard deviation of 10^{-4}. The histograms showing the relative errors can be seen in Figure 3b, and mean and maximum relative errors can be seen in the top part of Table 2.

For the same experiments with noisy data, we have implemented and run a Gauss-Newton optimizer, [27], to do local minimization of the sum of squared residuals. Over all measurements classified as inliers, $\sum \left(d_{ij} - ||\mathbf{r}_i - \mathbf{s}_j||_2\right)^2$ was locally minimized. To initialize the local optimizer, the calculated receiver and transmitter positions from the RANSAC-like solver were used. The histograms of

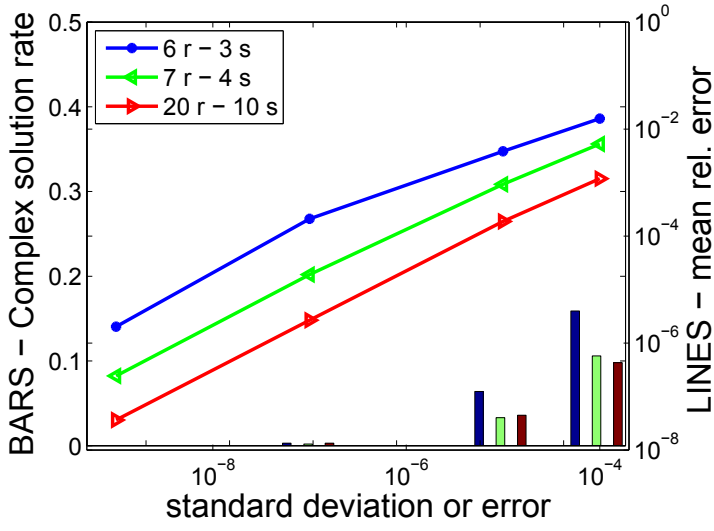
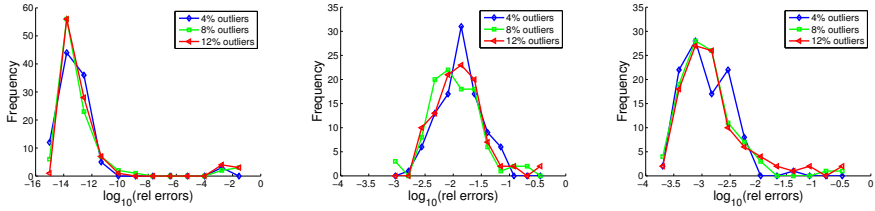


Figure 2: Relative errors for 1000 simulations each plotted against additive Gaussian noise, with different number of receivers (r) and transmitters (s) with $p = 2$. Box plot shows the percentage of complex solutions.

the relative errors can be seen in Figure 3c, and the maximum and mean relative errors can be seen in the bottom part of Table 2.

To have a basis for comparison to the RANSAC-like algorithm, we have implemented an iterated l_1-optimizer, [28], to minimize $\sum \left| d_{ij} - ||\mathbf{r}_i - \mathbf{s}_j|| \right|$ over all measurements, including outliers.We let the optimizer run on the same 100 experiments as above, but with random initializations of $\mathbf{r}_i$ and $\mathbf{s}_j$, randomized over the domain of the ground truth. The optimizer runs until approximate local convergence, and then starts with another random guess. This process continues until the time spent on the optimizer exceeds the time spent on the RANSAC-like solver for the same experiment. The best optimized solution is then saved. Over all the 100 experiments, the average number of random initializations optimized for each experiment was 13 runs. The results can be seen in Figure 4.

The failure rate for the RANSAC-algorithm is 0% in all cases. Theoretically, 300 runs of the main loop in the RANSAC-algorithm might not be enough to find a minimal case that could be used to trilaterate all transmitters, or alternatively, this might not be possible at all for the measurement matrix D in this



(a) (b) (c)

Figure 3: Histograms of relative errors for 100 simulations using the RANSAC-solver, for different number of outliers in measurements $D = [d_{ij}]$. 10 receiver and 15 transmitters were used. Measurements also has 1% missing data. (a) is without noise, (b) is with additive Gaussian noise with standard deviation 10^{-4}, and (c) is the same experiments as in (b) but with residuals $d_{ij} - ||\mathbf{r}_i - \mathbf{s}_j||_2$ locally minimized with Gauss-Newton.

given constellation, due to having too many outliers and missing data in an unfortunate constellation. For example, if all measurements in a column are missing data, one cannot hope to reconstruct the transmitter position corresponding to that column. However, a partial reconstruction of the node positions can be done in this case.

3.2 Real Data

Using seven Shure SV100 microphones and four Roxcore portable speakers, all connected to a Fast Track Ultra 8R sound card in an indoor environment, see Figure 5 (a), TOA measurements were obtained by matching sounds from different speakers to sound flanks recorded from different microphones. Another experiment was conducted, this time using T-bone MM1 measurement microphones, and new microphone and speaker positions. See Fig. 5 (b). The sounds in both experiments were separated temporally so that the matching of which sound came from which speaker could be done. Matching was done using the beginning of emitted sounds, thus ignoring reflections as there exist a direct path between speakers and microphones. Microphones were placed on a table, i.e. a plane, and speakers throughout the room, so that $p = 2$. The signal-to-noise ratio varied between 10-40 dB. A sample signal of a sound onset can be seen in Fig. 6.

To have a basis for comparison, a computer vision reconstruction of micro-

Outliers	Erroneously classified	False outlier	False inlier	Noise
4%	0.040%	0%	0.040%	No
8%	0.073%	0.013%	0.060%	No
12%	0.113%	0.020%	0.093%	No
4%	1.04%	0.980%	0.060%	Yes
8%	1.28%	1.17%	0.107%	Yes
12%	1.19%	0.980%	0.213%	Yes

Table 1: Over 100 experiments, the average % outliers in the measurements d_{ij} are tabulated against average of % i) measurements erroneously classified as inliers when outliers or vice versa, ii) measurements erroneously classified as outlier, iii) measurements erroneously classified as inlier, and iv) indicator if measurements have additive Gaussian noise of standard deviation 10^{-4}. All experiments also has an average 1% missing data.

phones and speakers were made, based on cell phone images of the experimental setup. The computer vision solution is then up to rotation, translation and scale. By running the extended algorithm, see Section 2.4, for $p = 2$ and aligning the solutions with the computer vision solutions, we got a root mean square error of 2.35 cm and 3.95 cm for the experiments using the Shure SV100 microphones and T-bone MM1 microphones respectively. See Figures 5 for a visualization.

4 Discussion

Looking at Figure 1, the results of the minimal solver of Algorithm 1 shows that the algorithm has good numerical performance. It is also fast, making it well suited for the RANSAC-like algorithm in Section 2.5.

The RANSAC-like algorithm shows good reconstruction results over all shown rate of outliers, seen in Figure 3a. The worst case scenario have relative error below 3%. As measurements are not corrupted by noise here, an empirical upper bound histogram for the reconstruction can be seen in Figure 1 for $p = 2$. The main deviation from this, the bump around the relative error $10^{-2.5}$, can be explained by one or more transmitters trilaterated from the initial 6 receivers has used one or more outliers from the data, that is similar to the correct measurement that should have been there. Only six measurements are available to trilaterate each new transmitter, whereas the minimal case is three measurements. Thus, there are

Outliers	Max relative error	Mean relative error	optimized
4%	7.71%	1.77%	No
8%	23.2%	1.88%	No
12%	35.3%	2.35%	No
4%	3.36%	0.20%	Yes
8%	23.4%	0.56%	Yes
12%	31.2%	0.99%	Yes

Table 2: Over 100 experiments with additive Gaussian noise of standard deviation 10^{-4}, the average % outliers in the measurements d_{ij} are tabulated against the relative errors of the reconstruction of receiver and transmitter positions, in % of i) max relative error , ii) mean relative error, iii) and iii) indicator if receiver and transmitter positions have been locally nonlinearly optimized. All experiments also has an average 1% missing data.

very few measurements at this stage to use in the internal trilateration RANSAC-approach. Furthermore, the final check when all receiver positions are available, counting the number of inliers of the 15 measurements belonging to the erroneous transmitter, does not tell us to retrilaterate the transmitter, as there is still a fair amount of inliers as the transmitter position is still fairly good.

Looking at the top part of Table 1, the erroneously classified measurements are few. Furthermore, some erroneous inliers are to be expected, as the outlier measurements are roughly randomized in the domain of measurement amplitude. Thus, we expect some of the outliers to be close to the correct value, and should perhaps not be said to be an outlier at all, but rather noisy measurements. The erroneous outliers again mostly coincide with cases where one or more transmitters have been reconstructed with errors like in the discussion above.

In the presence of outliers, moderate noise and missing data simultaneously, the RANSAC-like algorithm show fairly good reconstructions, as can be seen in Fiugre 3b. Looking at the lower part of Table 1, the erroneously classified measurements are still few. The measurements erroneously classified as outliers has increased somewhat in comparison with the noiseless case. This is to be expected as the minimal solvers for both the initial subset of 6 receivers and 3 transmitters and for the trilaterations only use minimal number of measurements, i.e. the solution fits the selected subset of measurements exactly, even though they are noisy.

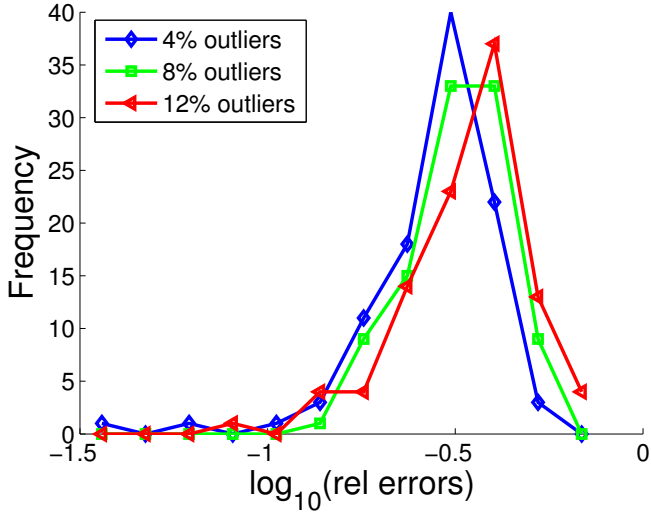


Figure 4: Histograms of relative errors for l_1 locally optimized solutions for 100 experiments, with several random initializations for each experiment. The time spent optimizing on each experiment is slightly exceeding the time spent on the RANSAC-like solver.

After locally minimizing the sum of squared residuals of the measurements classified as inliers, $\sum \left(d_{ij} - ||\mathbf{r}_i - \mathbf{s}_j||_2\right)^2$, a fair improvement of the reconstructions is made, see Figure 3c. The improvement can be explained by that the RANSAC-solver does not give a formally optimal solution, as it in each iteration fits receiver and transmitter positions only to a minimal subset of data. Figure 3b-3c show that the RANSAC-solver gives a good initial solution and a good classification of inliers, so that further local optimization can be done.

Comparing Figure 4 and 3b, we can see that the RANSAC-like algorithm performs much better than the best local l_1 optimization solution. Although the l_1 optimizer should be fairly robust to a few outliers, it can not reliably handle many outliers and the many local minima of the target function $\sum \left|d_{ij} - ||\mathbf{r}_i - \mathbf{s}_j||_2\right|$.

All experiments for the RANSAC-like algorithm uses 1% missing measurements. This shows that the method can handle missing data, but the real problem at hand is to identify outliers in the data and make a good receiver and transmitter position reconstruction.

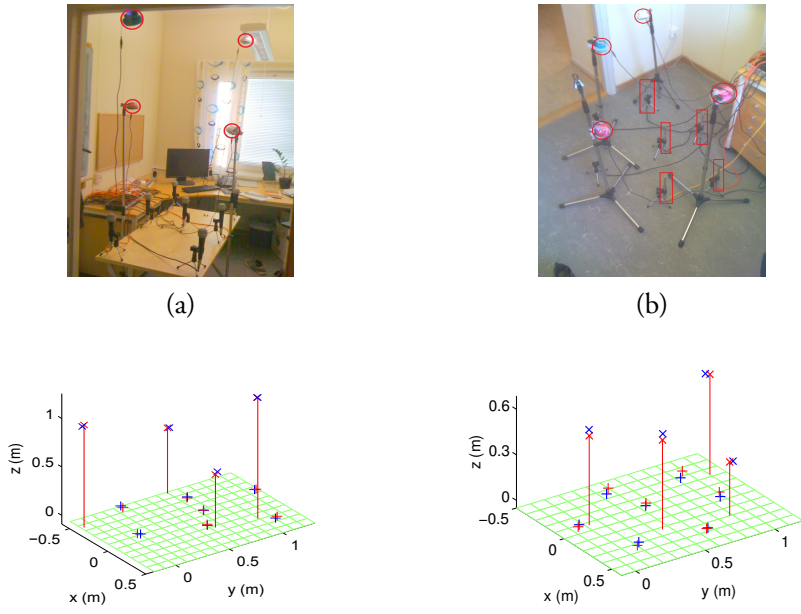


Figure 5: (a-b) Two experiments with speakers and microphones. Speakers are emphasized with circles and microphones with rectangles. Speakers are facing the microphone constellations. (b-c) The reconstructed solution using the algorithm (red) versus a reconstruction done with computer vision techniques (blue). Speakers are x and microphones are +.

5 Conclusions

In this paper we have solved the previously unsolved TOA calibration problem when receivers and transmitters are in different dimensional affine subspaces, for general dimensions. The primary interesting cases are when the lower dimensional subspace is on a line or in a plane, whereas higher dimensional solutions is for now of theoretical interest. The difference in dimensionality problem is an important degenerate case for previous papers focusing on the TOA calibration problem, and also for the TDOA calibration problem where a stratified approach of first determining offsets and then solving the TOA problem is a common approach. We solve the minimal case which is shown to be $1 + p + p(p + 1)/2$ receivers inhabiting a subspace of dimension p, and $p + 1$ transmitters inhabiting

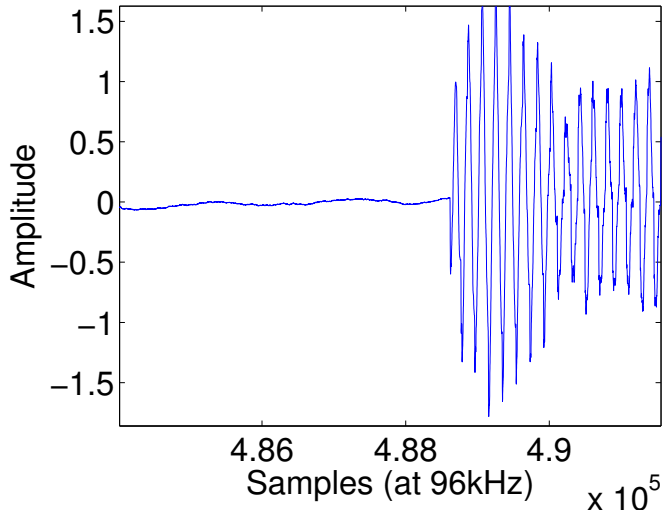


Figure 6: Typical microphone signal onset, taken from experiment shown in Fig. 5 (b)

a higher dimensional subspace. Only one solution exists up to gauge freedom and mirroring of transmitters in the subspace. We determine the degenerate cases and they are partially characterized. As a by-product we show how to solve the non-minimal TOA calibration problem for same dimensional case in general dimensions with linear techniques. The minimal algorithm is extended in a straightforward manner to be able to handle more than a minimal configuration of receivers and transmitters, and can serve as an initial estimate for non-linear optimization. A RANSAC-like algorithm has also been developed to show the feasibility of using the minimal algorithm for robust estimation and outlier removal in the TOA node self-calibration problem.

Simulated experiments support the feasibility of the algorithms. The minimal solver is numerically stable and the extended solver for overdetermined cases are tolerant to noise. Experiments on the RANSAC-like algorithm show small relative errors together with a strong ability to correctly classify measurements as outliers or inliers. Real indoor experiments using microphones and speakers in an acoustic sensor network support that the extension of the minimal algorithm can be used with good results to determine microphone and speaker positions.

References

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

[2] M. Cobos, A. Marti, and J.J. Lopez, "A modified srp-phat functional for robust real-time sound source localization with scalable spatial sampling," *Signal Processing Letters, IEEE Transactions on*, vol. 18, no. 1, pp. 71 –74, jan. 2011.

[3] Hoang Do, H.F. Silverman, and Ying Yu, "A real-time srp-phat source location implementation using stochastic region contraction(src) on a large-aperture microphone array," in *Proc. of International Conference on Acoustics, Speech, and Signal Processing*, 2007.

[4] Yosef Rockah and Peter M. Schultheiss, "Array shape calibration using sources in unknown locations - i: Far-field sources.," *Acoustics, Speech, and Signal Processing, IEEE Transactions on*, vol. ASSP-35, no. 3, pp. 286–299, 1987.

[5] M. Sun, L. Yang, and D. K. C. Ho, "Efficient joint source and sensor localization in closed-form," *IEEE Signal Processing Letters*, vol. 19, no. 7, pp. 399–402, 2012.

[6] L. Yang and K. C. Ho, "An approximately efficient tdoa localization algorithm in closed-form for locating multiple disjoint sources with erroneous sensor positions," *Signal Processing, IEEE Transactions on*, vol. 57, no. 12, pp. 4598–4615, 2009.

[7] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[8] Y. Kuang and K. Åström, "Stratified sensor network self-calibration from tdoa measurements," in *Proc. of the European Signal Processing Conference*, 2013.

[9] Leonard Asimow and Ben Roth, "The rigidity of graphs, ii," *Journal of Mathematical Analysis and Applications*, vol. 68, no. 1, pp. 171–190, 1979.

[10] T. Eren, OK Goldenberg, W. Whiteley, Y.R. Yang, A.S. Morse, BDO Anderson, and PN Belhumeur, "Rigidity, computation, and randomization in network localization," in *Proc. of Conference of the IEEE Communications Society*, 2004.

[11] E.D. Bolker and B. Roth, "When is a bipartite graph a rigid framework," *Pacific Journal of Mathematics*, vol. 90, no. 1, pp. 27–44, 1980.

[12] S. T. Birchfield and A. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 5, pp. 1025–1034, 2005.

[13] D. Niculescu and B. Nath, "Ad hoc positioning system (aps)," in *Proc. of Global Telecommunications Conference*, 2001.

[14] E. Elnahrawy, Xl. Li, and R. Martin, "The limits of localization using signal strength," in *Proc. of Sensor and Ad Hoc Communications and Networks*, 2004.

[15] V. C. Raykar, I. V. Kozintsev, and R. Lienhart, "Position calibration of microphones and loudspeakers in distributed computing platforms," *Speech and Audio Processing, IEEE Transactions on*, vol. 13, no. 1, pp. 70–83, 2005.

[16] M. Crocco, A. Del Bue, and V. Murino, "A bilinear approach to the position self-calibration of multiple sensors," *Signal Processing, IEEE Transactions on*, vol. 60, no. 2, pp. 660–673, 2012.

[17] J.C. Chen, R.E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *Signal Processing, IEEE Transactions on*, vol. 50, no. 8, pp. 1843–1854, 2002.

[18] S. Thrun, "Affine structure from sound," in *Proc. of Conference on Neural Information Processing Systems*, 2005.

[19] Y. Kuang, E. Ask, S. Burgess, and K. Åström, "Understanding toa and tdoa network calibration using far field approximation as initial estimate," in *Proc. of International Conference on Pattern Recognition Applications and Methods*, 2012.

[20] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, APR 2005.

[21] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2013.

[22] R. Biswas and S. Thrun, "A passive approach to sensor network localization," in *Proc. of International Conference on Intelligent Robots and Systems*, 2004.

[23] J. Wendeberg, F. Hoflinger, C. Schindelhauer, and L. Reindl, "Calibration-free tdoa self-localisation," *Journal of Location Based Services*, vol. 7, no. 2, pp. 121–144, 2013.

[24] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[25] K.S. Arun, T.S. Huang, and S.D. Blostein, "Least-squares fitting of two 3-d point sets," *Pattern analysis and machine intelligence, IEEE Transactions on*, vol. PAMI-9, no. 5, pp. 698–700, 1987.

[26] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2004, Second Edition.

[27] L Böiers, *Mathematical Methods of Optimization*, Studentlitteratur, 2010.

[28] E. J. Schlossmacher, "An iterative technique for absolute deviations curve fitting," *Journal of the American Statistical Association*, vol. 68, no. 344, pp. 857–859, 1973.

B

Paper B

Understanding TOA and TDOA Network Calibration using Far Field Approximation as Initial Estimate

Yubin Kuang, Erik Ask, Simon Burgess and Kalle Åström

Centre for Mathematical Sciences, Lund University, Lund, Sweden

Abstract

This paper presents a study of the so called far field approximation to the problem of determining both the direction to a number of transmittors and the relative motion of a single antenna using relative distance measurements. The same problem is present in calibration of microphone and wifi-transmittor arrays. In the far field approximation we assume that the relative motion of the antenna is small in comparison to the distances to the base stations. The problem can be solved uniquely with at least three motions of the antenna and at least six real or virtual transmittors. The failure modes of the problem is determined to be (i) when the antenna motion is planar or (ii) when the transmittor directions lie on a cone. We also study to what extent the solution can be obtained in these degenerate configurations. The solution algorithm for the minimal case can be extended to the overdetermined case in a straightforward manner. We also implement and test algorithms for non-linear optimization of the residuals. In experiments we explore how sensitive the calibration is with respect to different degrees of far field approximations of the transmittors and with respect to noise in the data.

Key words: TOA, TDOA, calibration, sensor networks, self-localization.

1 Introduction

Navigation covers a broad application area ranging from traditional needs in the terrestrial, aerial and naval transport sectors to personal objectives of finding your way to school if you are visually impaired, to the nearest fire exit in case of an emergency, or to specific goods in your local supermarket. Many potential applications are however presently hindered by performance limitations of existing positioning techniques and navigation systems.

Radio based positioning rely on either signal strength, direction of arrival (DOA) or time-based information such as time of arrival (TOA) or time differences of arrival (TDOA), or a combination thereof.

The identical mathematical problem occurs also in microphone arrays for audio sensing. Using multiple microphones it is possible to locate a particular sound-source and using beamforming to enhance sound quality of the speaker.

Although TOA and TDOA problems have been studied extensively in the literature in the form of localization of e.g. a sound source using a calibrated detector array, the problem of calibration of a sensor array using only measurement, i.e. the initialization problem for sensor network calibration, has received much less attention. One technique used for sensor network calibration is to manually measure the inter-distance between pairs of microphones and use multidimensional scaling to compute microphone locations, [1]. Another option is to use GPS, [2], or to use additional transmittors (radio or audio), close to each receiver, [3–5]. Sensor network calibration is treated in [6]. In [7] it is shown how to estimate additional microphones, once an initial estimate of the position of some microphones are known. In [8] the far field approximation is used to initialize the calibration of sensor networks. However the experiments and theory was only tested for the planar case and no study of the failure modes were given. Initialization of TOA networks has been studied in [9], where solutions to the minimal case of three transmittors and three receivers in the plane is given. The minimal case in 3D is determined to be four receivers and six transmittors for TOA, but this is not solved. Initialization of TDOA networks is studied in [10], where solutions were give to two non-minimal cases of ten transmittors and five receivers, whereas the minimal solution for far field approximation in this paper are six transmittors and four receivers.

In this paper we study far field approximation as an initialization to the calibration problem. We use a similar factorization as [8] but in three dimensions, and show that far field approximation is at least four measurement positions, i.e.

three motions, and measurements to at least six real or virtual transmittors. In this paper we describe the failure modes of the algorithm and show what can be done when such configurations are present. We further propose two optimization strategies for more thorough calibration and evaluate them in regards to accuracy and convergence rate. Several test cases are simulated in which we validate far field approximation, accuracy of the proposed algorithms, optimization schemes and performance under noisy measurements.

2 Determining Pose

In the following treatment, we make no difference between real and virtual transmittors or base stations. Assume that the base station is stationary at position $b = \begin{pmatrix} b_x & b_y & b_z \end{pmatrix}$ and that the antenna is at position $z = \begin{pmatrix} z_x & z_y & z_z \end{pmatrix}$. By measuring the signal with known base band frequency one obtains a complex constant, whos phase depends on the distance $d = |b - z|$ between the antenna and the base station.

By tracking the phase during small relative motions of the antenna, it is feasible to determine the relative distance $d_{rel}(t) = d(t) + \tilde{C}$, where $\tilde{C}$ is an unknown constant for each base station. This is the so called TDOA setup. Furthermore if during measurements the relative motion is small in comparison with the distance d between the antenna z and the base station b it is reasonable to approximate the distance $d = |b-z| \approx |b-z_0| + (z-z_0)^T\mathbf{n} = z^T\mathbf{n} + \underbrace{(|b-z_0| - z_0^T\mathbf{n})}_{\bar{C}}$. Here z_0 is the initial position of the antenna and $\mathbf{n}$ is the direction from the base station towards the antenna, now assumed to be constant with unit length. By setting $C = \tilde{C} + \bar{C}$ one obtains the far field approximation

$$d_{rel}(\mathbf{n}, z) \approx z^T\mathbf{n} + C.$$

In this paper we are interested in the following far field time difference of arrival (FFTDOA) type problem that arise from this approximate relative distance measurement.

Problem 2.1. Given measurements $D_{i,j},\ i = 1, \ldots, m$ *and* $j = 1, \ldots, k$ from the antenna at m different positions to k base stations, determine both the positions $z_1, \ldots, z_m$ of the antenna during the relative motion and the directions $\mathbf{n}_1 \ldots \mathbf{n}_k$ from the base stations so that

$$D_{i,j} = z_i^T\mathbf{n}_j - C_j,$$

$$||\mathbf{n}_j||_2 = 1$$

where C_j is a constant distance offset for each base station.

Lemma 2.1. *A problem with m measurements to k base stations with unknown constant C_j can without loss of generality be converted to a problem with $m-1$ measurements to k base stations with known constant.*

Proof. Note that because of the unknown constant C_j the problem does not change in character by modification $\bar{D}_{i,j} = D_{i,j} - K_j$. For simplicity we set $\bar{D}_{i,j} = D_{i,j} - D_{1,j}$. By also setting $z_1 = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^T$, we get $C_j = 0$. This is equivalent to choosing the origin of the unknown coordinate system to the first point. □

For simplicity we will in the sequel assume that $C_j = 0$ and assume that the one measurement has already been used to resolve the ambiguity. Denote by D the matrix after removing that said point. This converts the FFTDOA problem into a FFTOA problem, i.e.

Problem 2.2. Given measurements $D_{i,j},\ i = 1,\ldots,m, j = 1,\ldots,k$ from the antenna at m different positions to k base stations, determine both the both the positions z_i of the antenna during the relative motion and the direction from the base stations $\mathbf{n}_j$ so that

$$D_{i,j} = z_i^T \mathbf{n}_j.$$
$$||\mathbf{n}_j||_2 = 1$$

Lemma 2.2. *The matrix D with elements $D_{i,j}$ is of rank at most 3.*

Proof. The measurement equations are $D_{i,j} = z_i^T \mathbf{n}_j$. By setting

$$Z = \begin{pmatrix} z_1^T \\ z_2^T \\ \vdots \\ z_m^T \end{pmatrix}$$

and

$$N = \begin{pmatrix} n_1 & n_2 & \ldots & n_k \end{pmatrix}$$

we see that $D = ZN$. Both Z and N have at most rank 3, therfore the same holds for D. □

Assuming that k and m are large enough and assuming that the motion z_i and the base stations $\mathbf{n}_j$ are in general enough constellation the matrix D will have rank 3. If so it is possible to reconstruct both Z and N up to an unknown linear transformation. This can be done using singular value decomposition, $D = USV^T$. Even with noisy measurements, the closest rank 3 approximation in the L_2 norm can be found using the first 3 columns of U and V. By setting $\tilde{Z} = U_3$ and $\tilde{N} = S_3 V_3^T$ we get all possible solutions by $N = A\tilde{N}$, with A a general full rank 3×3 matrix. Changing A corresponds to rotating, affinely stretching and possibly mirroring the coordinate system. The true reconstruction also fulfills $\mathbf{n}_j^T \mathbf{n}_j = 1$, which gives constraints on A of type

$$\mathbf{n}_j^T A^T A \mathbf{n}_j = 1,$$

which after substitution $B = A^T A$ becomes linear

$$\mathbf{n}_j^T B \mathbf{n}_j = 1$$

in the unknown elements of B. Since symmetric 3×3 matrices have 6 degrees of freedom we need at least 6 base stations to determine the matrix uniquely. Once B has been determined A can be determined by Cholesky factorization. This gives the transformation A up to an unknown rotation and possible mirroring of the coordinate system. We summarize the above in the following theorem.

Theorem 2.1. *The minimal case for reconstructing m positions z_i and k orientations $\mathbf{n}_j$ from relative distance measurements $D_{i,j}$ as formulated in Problem 2.2 is $m = 4$ and $k = 6$.*

Accordingly, we present Algorithm 1 for the minimal case of the problem. Note that using minimal information $m = 4$ and $k = 6$ results in estimates that fulfill the measurements exactly (up to machine precision) even if the measurements are disturbed by noise.

2.1 Failure Modes of the Algorithm

It is interesting and enlightening to know the failure modes of the algorithm. This is captured by the following theorem.

Theorem 2.2. *The minimal case for reconstructing m orientations $\mathbf{n}_j$ and k positions z_i from relative distance measurements $D_{i,j}$ as formulated in Problem 2.2 is for $m = 4$*

Algorithm 1 Minimal algorithm for the FFTDOA problem

Input: Measurement matrix D of size 4×6.

1: Set $\bar{D}_{i,j} = D_{i,j} - D_{1,j}$
2: Remove the first row of $\bar{D}$
3: Calculate a singular value decomposition $\bar{D} = USV^T$.
4: Set $\tilde{Z}$ to first 3 columns of U and $\tilde{N}$ to first three columns of SV^T.
5: Solve for the six unknowns in the symmetric matrix B using the 6 linear constraints $\tilde{n}_j^T B \tilde{n}_j = 1$.
6: Calculate A by Cholesky factorization of B, so that $A^T A = B$.
7: Transform motion according to $Z = \tilde{Z}A^{-1}$ and structure according to $N = A\tilde{N}$.

and $k = 6$. As long as the orientations $\mathbf{n}_j$ do not lie on a common quadratic cone $\mathbf{n}_j^T \Omega \mathbf{n}_j = 0$ and the measurement positions z_i do not lie on a plane, there will not be more than one solution to the problem of determining both structure $\mathbf{n}_j$ and motion z_i up to an unknown translation, orientation and reflection of the coordinate system.

Proof. The algorithm can fail if the measurement matrix D has rank 2 or lower. This could e.g. happen if either all measurement positions z_i lie in a plane or if all directions $\mathbf{n}_j$ lie in a plane (or both). The algorithm can also fail if there are two solutions to the matrix B in $\mathbf{n}_j^T B \mathbf{n}_j = 1$. But then the difference $\Omega = B_1 - B$ of these two solutions is a three by three matrix for which

$$\mathbf{n}_j^T \Omega \mathbf{n}_j = 0,$$

which in turn implies that the directions $\mathbf{n}_j$ lie on a common conic as represented by the matrix Ω. □

Yet another type of failure mode of the algorithm is if the data is corrupted by noise or far field approximation is not vaild, so that the matrix B obtained is not positive definite. Then the algorithms fails because there is no Cholesky factorization of B into $A^T A$. If B is unique, there are no real solution to the problem in this case.

2.2 Analysis of Failure Modes

If the rank of the matrix D is 2, this could be because the points z_i lie on a plane or that $\mathbf{n}_j$ lie on a plane.

In this case of coplanar z_i it is still possible to estimate the planar coordinates $Z = U_2 A$ and $N = AS_2V_2^T$ up to an unknown 2×2 matrix A representing a choice of affine coordinate system. Here we do get inequality constraints that

$$\left| A \begin{pmatrix} n_{j,x} \\ n_{j,y} \end{pmatrix} \right| \leq 1.$$

Each such A is a potential solution. It is possible to extend with a third coordinate in the normal direction according to

$$n_{j,z} = \pm\sqrt{1 - n_{j,x}^2 - n_{j,y}^2}.$$

Another possibility is that the directions $\mathbf{n}_j$ lie on a plane. In this case it is possible to reconstruct two of the coordinates for both the positions z_i and the directions $\mathbf{n}_j$. Since the normals are assumed to lie in a plane, we can exploit the equality constraints $\mathbf{n}_j^T A^T A \mathbf{n}_j = 1$ similar to the rank 3 case. In this particular case we only need three directions $\mathbf{n}_j$, i.e. the minimal case is for $m = 3$ and $k = 3$. This gives the full reconstruction of both points and directions up to an unknown choice of Euclidean coordinate system and unknown choice of z-coordinate for the points z_i.

If the rank is 1, this could be because the directions are parallel. In this case. Similar to the discussions above we can obtain one of the coordinates of the positions z_i, but this is trivial since the measurements $D_{i,j}$ are such coordinates by definition.

If the rank is 1 because the points lie on a line, we obtain a one-parameter family of reconstructions based on $Z = U_1 a$ and $N = aS_1V_1^T$, where a is an unknown constant that has to fulfill $a \leq 1/l$, where $l = max_j |S_1 V_{1,j}|$. For each such a it is possible to extend the directions $\mathbf{n}_j$ so that they have length one, but there are several such choices.

2.3 Overdetermined Cases

When more measurements are available than the minimal case discussed in the previous section, we need to solve an overdetermined system in least-square sense or with robust error measures e.g. L_1-norm. Here we focus on the following least-square formulation for the pose problem:

Problem 2.3. Given measurements $D_{i,j},\ i = 1, \ldots,\ m$ *and* $j = 1, \ldots, k$ from the antenna at m different positions to k base stations, determine both the relative motion of the antenna $\mathbf{z}_i$ and the direction to the base stations $\mathbf{n}_j$ so that

$$\begin{aligned} \min_{Z,N} \quad & ||D - Z^T N||^2_{Frob} \\ s.t. \quad & ||\mathbf{n}_j||_2 = 1,\ j = 1, \ldots, k. \end{aligned} \tag{1}$$

where $||.||_{Frob}$ denotes the Frobenius norm.

For the over-determined cases, that is $m > 4$ and $k \geq 6$ or $m \geq 4$ and $k > 6$, it is possible to modify Algorithm 1 to obtain an efficient but not necessarily optimal algorithm that finds a reconstruction that fits the data quite good using the following three modifications (i) the best rank 3 approximation can still be found in step 4-5 using the singular value decomposition, (ii) the estimate of B in step 6 can be performed in a least squares sense and (iii) re-normalize the columns of N to length 1. This results in a reconstruction that differs from the measurements, but both steps are relatively robust to noise. The problem of B not being positive semi-definite can be attacked by non-linear optimization. Here we try to optimize A so that $\sum_{j=1}^{k} (\mathbf{n}_j^T A^T A \mathbf{n}_j - 1)^2$ is minimized. This can be achieved e.g. by initializing with $A = I$ and then using non-linear optimization of the error function.

Clearly, we lose any guarantee on the optimality of the solution when we enforce the constraints as in step (iii). However, the solution can serve as a good initialization for subsequent optimization algorithms we present in this section. We discuss how to use alternating optimization and Levenberg-Marquardt algorithm (LMA) to obtain better solution. The first algorithm starts with an initial feasible solution for Z and N, and then it alternates between optimizing Z given N and vice versa. The latter is essentially a method combining Gauss-Newton algorithm and a gradient descent that improve the solution locally. For both methods, we need to treat the constraints on the direction vectors properly to ensure convergence.

2.3.1 Alternating Optimization

In order to find the local minima of Problem 2.3, we can use a coordinate descent scheme. Specifically, we would like to iteratively optimize the cost function in Problem 2.3 with respect to Z given N, and then find the optimal feasible N with fixed Z. If we initialize N such that it satisfies the norm constraints, we can easily see that the alternating procedure is converging (Algorithm 2).

Algorithm 2 Alternating optimization for the FFUTOA problem

Input: Measurement matrix D with $m > 4$ and $k \geq 6$ or $m \geq 4$ and $k > 6$,

1: Construct $\bar{D}$ and initialize Z and N as in Algorithm 1
2: Fix N , find optimal Z
3: Fix Z, solve the constrained minimization for each $\mathbf{n}_j$, $j = 1, ..., k$
4: Repeat (2) and (3) until convergence or predefined number of iterations is reached

To enable the alternating optimization, we need to solve two separate optimization problems. The first one is to find the optimal Z given N. This is the classic least squares problem and is known to be convex and can be solved efficiently. On the other hand, solving for optimal $\mathbf{n}_j$ given Z is not always convex due to the additional constraints on the $\mathbf{n}_j$'s. In this case, we seek the local minima for each $\mathbf{n}_j$ as a constrained minimization problem. We solve the small constrained problems (3 variables each) independently with interior point method. Alternatively, we can solve the constrained optimization as solving polynomial equations. This can be related to the fact that for a given Z, level sets of the cost function with respect to $\mathbf{n}_j$ are surfaces of a ellipsoid in $\mathbb{R}^3$ (the centers are in this case the solution from singular value decomposition). The norm 1 constraints on $\mathbf{n}_j$ geometrically means that the feasible solutions lie on the unit sphere centered at origin. Therefore, the optimal solution of $\mathbf{n}_j$ is one of the points that the ellipsoid is tangent to the unit sphere, which can be found by solving polynomial equations. While there could exist multiple solutions, we can choose the one with minimum euclidean distances to the center of the ellipsoid. Unlike interior point solver, we always find the global optimum. However, in practice, we found that in the alternating procedure, interior point method and polynomial solving give similar performance.

2.3.2 Levenberg-Marquardt Algorithm

It is well-known that alternating optimization as a coordinate descent scheme converges slowly in practice. Alternatively, we can solve the minimization problem by iteratively finding the best descent direction for N and Z simultaneously. The difficulty here is again the constraints on the direction vectors $\mathbf{n}_j$. The key idea here is to re-parameterize the orientation vectors. Given a direction vector $\mathbf{n}$ having unit length, any direction vectors can be represented by $\mathbf{n} \cdot exp(S)$, where S

Algorithm 3 Levenberg-Marquardt optimization for the FFUTOA problem

Input: Measurement matrix D (over-determined), initialize $\mathbf{y}$ and construct $\bar{\mathbf{d}}$ as in Algorithm 1,

1: Compute the Jacobian of $\bar{\mathbf{d}}$ with respect to $\mathbf{y}$, $J = (\frac{\partial \bar{D}_{11}}{\partial \mathbf{y}}, \ldots, \frac{\partial \bar{D}_{ij}}{\partial \mathbf{y}}, \ldots, \frac{\partial \bar{D}_{(m-1)k}}{\partial \mathbf{y}})$

2: Calculate $\Delta\mathbf{y} = (J^T J + \lambda \cdot \mathrm{diag}(J^T J))^{-1} J^T \Delta\bar{\mathbf{d}}$, where $\Delta\bar{\mathbf{d}}$ is the residue and λ a damping factor.

3: $\mathbf{y} = \mathbf{y} + \Delta\mathbf{y}$

4: repeat (1),(2) and (3) until convergence or predefined number of iterations is reached

is a 3×3 skew-symmetric matrix. This is due to the fact that the exponential map of any such matrix is a rotation matrix. In this case, if we use the (current) orientation $\mathbf{n}$ as axis direction, any local change of the orientation on the sphere can be easily parameterized via the exponential map. Therefore, the gradient of D_{ij} with respect to $\mathbf{n}_j$ can be expressed without any constraints. We can then construct the Jacobian for the Levenberg-Marquardt algorithm to compute the optimal descent direction. In the following, we use $\mathbf{y}$ to denote the vector formed by stacking variables in Z and N, $\bar{\mathbf{d}}$ is the vectorized version of $\bar{D}$ based on the ordering of g.

3 Experimental Validation

In this section, we present comprehensive experimental results for simulated data. We focus on the performance of the minimal solver, verification of the far field approximation as well as the comparisons between solvers for overdetermined cases.

3.1 Minimal Solver Accuracy

The numerical performance of the algorithm was evaluated by generating problems where the far field approximation is true and not degenerate. In essence this constitutes generating directions $\mathbf{n}_j$ and relative distance measurements $D_{i,j}$ and culling cases where the three largest singular values of the measurement matrix aren't above a threshold or the directions lie on a conic. The error is then evaluated as the average norm-difference of the estimated reciever positions. The

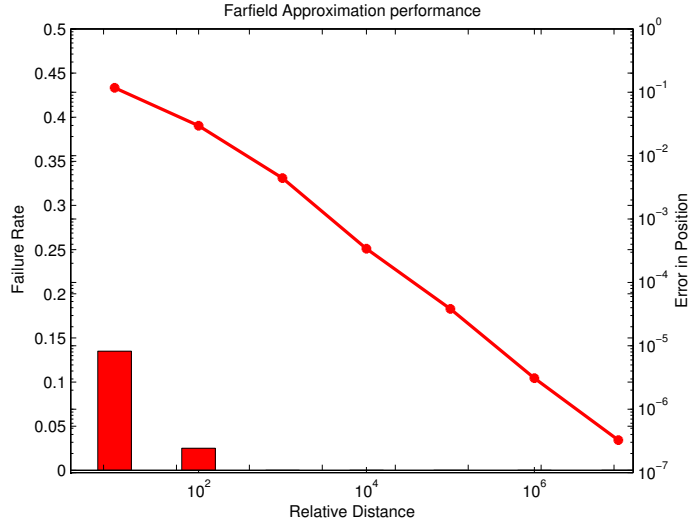


Figure 1: Performance on minimal case solver. Bars show failure rate (left y-axis) and curve shows the norm of the error in position as a function of distance (right y-axis). Note that bar height has linear scale.

reciever positions were selected as the corners of a tetrahedron with arc-length one. The average error of 10000 such tests was $6.8 \cdot 10^{-15}$, close to machine epsilon.

3.2 Far Field Approximation Accuracy

3.2.1 Minimal Case

To evaluate the performance of the assumption that senders can be viewed as having a single common direction to receivers, data was generated using 3D positions for both senders and receivers at different relative distances in-between receivers and senders to receivers. The constellation of receivers is again the tetrahedron and senders are randomly placed on a sphere surrounding it. A graph showing the error, as defined in section 3.1, as a function of radius of the sphere (that is relative distance), as well as the failure rate of the solver is shown in Figure 1. A failure constitutes a case in which the B matrix in algorithm 1 is not positive definite. As can be seen this is infrequent even at small relative distances in when

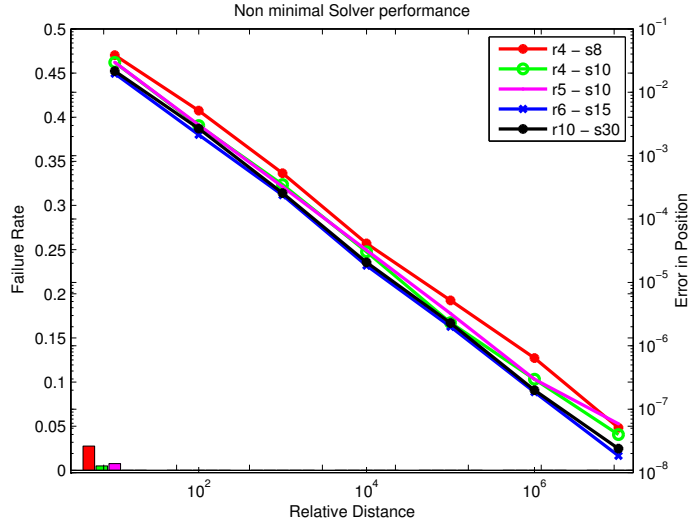


Figure 2: Performance on non-minimal cases. Bars show failure rates, line is error as a function of distance in loglog scale. Size of test-cases are noted in figure with rx-sy denoting x recievers and y senders. This plot is best viewed in color. Note the scale difference to the graph in Figure 1.

one would not expect a far field approximation to work. As can be expected the approximation gets better when the relative distance increases.

3.2.2 Initialization for Overdetermined Cases

As described in section 2.3, Algorithm 1 can with some modifications be used on overdetermined cases without guarantees on optimality of the solution. In these situations the solutions serves as an initial guess of some other optimization method. The additional information should however give some numerical stability and it is interesting to evaluate the algorithm for initial guess estimates in overdetermined cases. To do this the synthetic dataset is augmented with additional randomly placed senders and receivers. The four first receivers are again the tetrahedron and the rest are randomly uniformly distributed within the unit cube. Senders are again placed on a sphere around the receivers. Results for different problem sizes are shown in Figure 2. One immediately notices that the failure ratio drops, in many cases to zero. One can also see that adding more data will

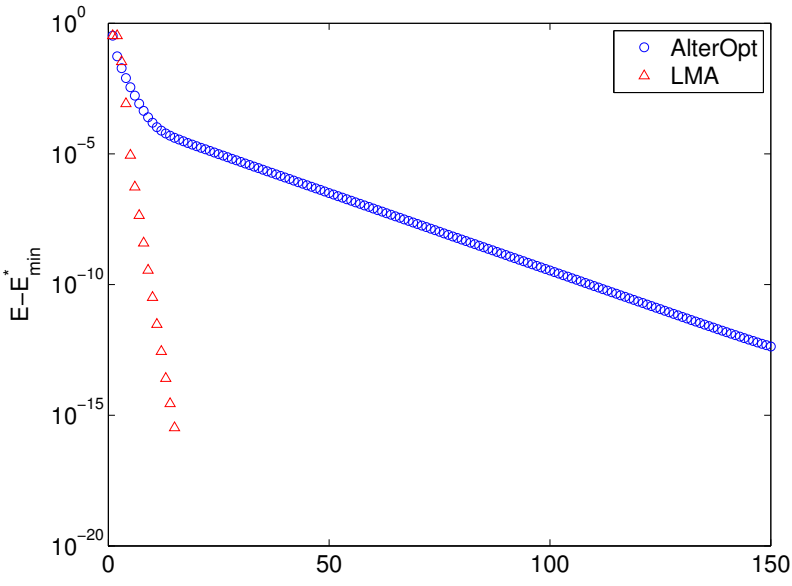


Figure 3: Convergence of alternating optimization and LMA on simulated TDOA measurements with gaussian white noise ($\sigma = 0.1$). Here $m = 10$ and $k = 10$.

(in general) result in smaller errors, for the cases shown here up to one order of magnitude smaller than a min case.

3.3 Overdetermined Cases

We also investigate the performance of the two schemes for over-determined cases. In all experiments below, we initialize both the alternating optimization and LMA based on the minimal solver modified for over-determined case. The simulated data is of a true far field approximation with gaussian white noise, i.e. measurements are simulated as $D_{i,j} = \mathbf{z}_i\mathbf{n}_j + \varepsilon_{i,j}$ where $\varepsilon_{i,j} \in N(0, \sigma)$ i.i.d. In Figure 3, we can see that alternating optimization and LMA all decrease the reconstruction errors compared to the minimal solver. On the other hand, from figure 3, LMA converges much faster than alternating scheme (20 vs. 150) and obtains relatively lower reconstruction errors. This verifies the superiority of LMA over coordinate descent. This observation is consistent over different m and k as well as a variety

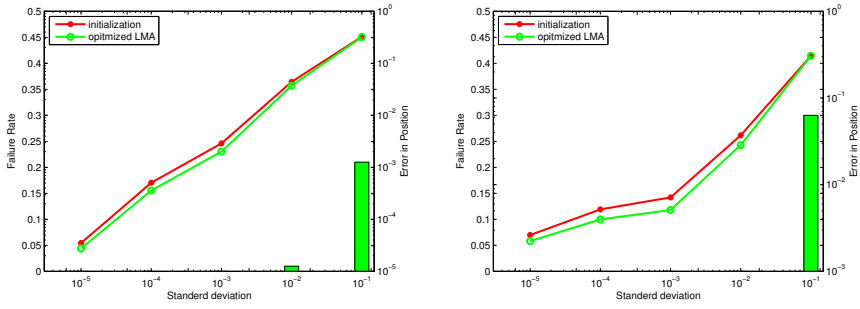


Figure 4: Performance on non-minimal cases with simulated TDOA measurements with gaussian white noise. The mean error in position of the receivers are plotted against the noise standard deviation. Here m=5 and k=10, and the relative distance to receivers and senders are 10^7 (left) and 10^2 (right). Failure rates for the initialization are also shown for completeness.

of noise levels. Note that here for all the experiments, we set the damping factor λ to 1.

It it also of interest to view the complete system when the measurements $D_{i,j}$ does not fulfill the far field approximation and when disturbed by noise. The relative distances of the simulated senders and receivers are set to 10^2 for a mediocre far field approximation and 10^7 for a good far field approximation. TDOA measurements $D_{i,j}$ are then simulated, perturbed with gaussian white noise. Figure 4 shows the results. The pictures show that the initialization method is fairly good, but in many cases the LMA brings down the position error. The system is also fairly robust to noise.

4 Conclusions

In this paper we study the far field approximation to the calibration of TDOA and TOA sensor networks. The far field approximation of the problem results in a factorization algorithm with constraints. The failure modes of the algorithm is studied and particular emphasis is made on what can be said when any of these failure conditions are met. The experimental validation gives a strong indication that a far field approximation is a feasible approach both for getting direct estimates as well as initial estimates for other solvers. Even considering that there are

cases when the algorithm fails, obtained solutions are good even at small relative distances. This validation is done on 3D problems and confirms findings in [8] where evaluation was done in 2D.

Further we analyze two optimization schemes and what difficulties may arise when employing them. Both of these schemes are experimentally evaluated and confirmed to successfully optimize the initial guess on a problem fulfilling the far field assumptions, although at quite different convergence rates. The faster of the two is also employed on cases when senders are given true locations and measurements are subject to noise with good results.

It would be interesting in future work to study to what extent it can be shown that the local optimum obtained to the problem can be proven to be global optimum. To integrate the solvers with robust norms is also worth studying to handle situations with outliers. It would also be interesting to verify the algorithms on real measured data and investigate the possibilities of using our algorithms in a RANSAC approach to remove potential outliers that may occur in real life settings.

Acknowledgements

The research leading to these results has received funding from the Swedish strategic research projects ELLIIT and ESSENSE, the Swedish research council project Polynomial Equations and the Swedish Strategic Foundation projects ENGROSS and Wearable Visual Systems.

References

[1] S. T. Birchfield and A. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *Speech and Audio Processing, IEEE transactions on*, vol. 13, no. 5, 2005.

[2] D. Niculescu and B. Nath, "Ad hoc positioning system (aps)," in *GLOBECOM-01*, 2001.

[3] E. Elnahrawy, Xl. Li, and R. Martin, "The limits of localization using signal strength," in *SECON-04*, 2004.

[4] V. C. Raykar, I. V. Kozintsev, and R. Lienhart, "Position calibration of microphones and loudspeakers in distributed computing platforms," *Speech and Audio Processing, IEEE transactions on*, vol. 13, no. 1, 2005.

[5] J. Sallai, G. Balogh, M. Maroti, and A. Ledeczi, "Acoustic ranging in resource-constrained sensor networks," in *eCOTS-04*, 2004.

[6] R. Biswas and S. Thrun, "A passive approach to sensor network localization," in *IROS 2004*, 2004.

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

[8] S. Thrun, "Affine structure from sound," in *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, Cambridge, MA, 2005, MIT Press.

[9] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, 2005.

[10] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

C

Paper C

Minimal Solvers for Unsynchronized TDOA Sensor Network Calibration

Simon Burgess[1], Yubin Kuang[1], Johannes Wendeberg[2], Kalle Åström[1] and Christian Schindelhauer[2]

[1] *Centre for Mathematical Sciences, Lund University, Lund, Sweden*
[2] *Department of Computer Science, University of Freiburg, Freiburg, Germany*

Abstract

We present two novel approaches for the problem of self-calibration of network nodes using only TDOA when both receivers and transmitters are unsynchronized. We consider the previously unsolved minimum problem of far field localization in three dimensions, which is to locate four receivers by the signals of nine unknown transmitters, for which we assume that they originate from far away. The first approach uses that the time differences between four receivers characterize an ellipsoid. The second approach uses linear algebra techniques on the matrix of unsynchronized TDOA measurements. This approach is easily extended to more than four receivers and nine transmitters. In extensive experiments, the algorithms are shown to be robust to moderate Gaussian measurement noise and the far field assumption is reasonable if the distance between transmitters and receivers is at least four times the distance between the receivers. In an indoor experiment using sound we reconstruct the microphone positions up to a mean error of 5 cm.

1 Introduction

In this paper we study the problem of node localization using only Unsynchronized Time difference Of Arrival (UTOA) measurements between nodes, where either receivers or transmitters are far away from the other group. The problem arises naturally in microphone arrays for audio sensing. Is it possible to calculate both multiple microphone positions as well as the timings and directions of the sound sources, if the microphones are unsynchronized, i.e. do not use the same clock, just from sounds emanating from far away at unknown locations and times? An example application could be to locate several cell phones just by environmental sounds, where cell phone positions and sound directions are to be recovered without synchronizing the phones first.

1.1 Related Work

Although time of arrival (TOA) and time difference of arrival (TDOA) problems have been studied extensively in the literature in the form of localization of e.g. a sound source using a calibrated array, see e.g. [1–4], the problem of calibration of a sensor array from only measurements, i.e. the node localization problem, has received less attention.

In [5] and refined in [6] a far field approximation was utilized to solve the TOA and TDOA case, with the minimal number of four receivers and six unsynchronized far field transmitters in 3D. Under the assumption that signals and receivers are distributed in the unit disk, the distance between receivers can be approximated by evaluation of the range of time differences [7–9] or by statistical analysis of their distribution [10, 11], although these approaches depend on the availability of a large number of signals. Calibration of TOA networks using only measurements has been studied in [12, 13], where solutions to the minimal cases of three transmitters and three receivers in the plane, or six transmitters and four receivers in 3D are given. Calibration of TDOA networks is studied in [14] and further improved upon in [15], where the non-minimal case of eight transmitters and five receivers is solved. In [16, 17] a TDOA setup is used for indoor navigation based on non-linear optimization, but the methods can get stuck in local minima and are dependent on initialization.

The problem of node localization using only UTOA measurements from unsynchronized receivers and transmitters in a far field setting has been considered in [18], however the approach requires at least five receivers, which is more

than the minimum case. Minimal algorithms are of importance in RANSAC schemes [19] to weed out outliers in noisy data which is a common problem in TOA/TDOA/UTOA applications. The problem has been addressed in a different manner estimating ellipse coefficients in [20], but no analysis of degenerate cases has been done and the algorithm is only described for the planar case.

In this paper we expand on previous work and propose two novel algorithms for parameter estimation of a receiver array, the Ellipsoid method in 3D and the Matrix Factorization method for UTOA measurements, that both consider the minimum case of four receivers and nine transmitters in three dimensions. We compare the methods on simulated and real data where we demonstrate their numerical stability. The methods are also evaluated on overdetermined cases using more than four receivers and nine transmitters.

2 Problem Setting

In the following treatment, we make no difference between real and virtual transmitters. Assume that the transmitters are stationary at position $\mathbf{b}_j \in \mathbb{R}^3$, $j = 1, \ldots, k$ and that the receivers are at positions $\mathbf{r}_i \in \mathbb{R}^3$, $i = 1, \ldots, m$. By measuring how long time the signals take to reach the receiver and knowing the speed of the signals, distances $\delta_{ij} = \|\mathbf{r}_i - \mathbf{b}_j\|$ can be measured, $\|\cdot\|$ denoting the Euclidean norm. These are TOA measurements.

When neither receivers or transmitters are synchronized, for instance external sound sources recorded on different computers, the measurements will be of the form $\delta_{ij} = \|\mathbf{r}_i - \mathbf{b}_j\| + f_i + \tilde{g}_j$ where $f_i, \tilde{g}_j$ are unknown offsets for receivers and transmitters respectively. We denote measurements of this kind Unsynchronized Time difference Of Arrival (UTOA) measurements. Furthermore, if the transmitters are so far from the receivers that a transmitter can be considered to have a common direction to the receivers, the measurements can be approximated by

$$\delta_{ij} = \|\mathbf{r}_i - \mathbf{b}_j\| + f_i + \tilde{g}_j \approx \|\mathbf{r}_1 - \mathbf{b}_j\| + (\mathbf{r}_i - \mathbf{r}_1)^T \mathbf{n}_j + f_i + \tilde{g}_j = \mathbf{r}_i^T \mathbf{n}_j + \bar{g}_j + f_i + \tilde{g}_j \quad (1)$$

where $\bar{g}_j = \|\mathbf{r}_1 - \mathbf{b}_j\| - \mathbf{r}_1^T \mathbf{n}_j$ and $\mathbf{n}_j$ is the direction of unit length from transmitter j to the receivers. By setting $g_j = \bar{g}_j + \tilde{g}_j$ we get the far field approximation

$$\delta_{ij} \approx \mathbf{r}_i^T \mathbf{n}_j + f_i + g_j.$$

When the approximation is good, we will call δ_{ij} Far Field UTOA (FFUTOA) measurements.

2.1 The FFUTOA Calibration Problem

We assume that (i) the speed of signals v is known, and thus all time measurements are transformed to distances by multiplication by v and (ii) receivers can distinguish which TOA signal comes from which sender. This can be done in practice by e.g. separating the signals temporally or by frequency.

Problem 2.1. Given mk FFUTOA measurements $\delta_{ij} \in \mathbb{R}, i = 1,\dots,m, j = 1,\dots,k$, taken from m receivers and k transmitters, estimate receiver positions $\mathbf{r}_i \in \mathbb{R}^3$, directions $\mathbf{n}_j \in \mathbb{R}^3$ from transmitter j to receivers, receiver and transmitter offsets $f_i \in \mathbb{R}$, $g_j \in \mathbb{R}$ such that

$$\delta_{ij} = \mathbf{r}_i^T \mathbf{n}_j + f_i + g_j \ , \quad \text{and} \quad \|\mathbf{n}_j\| = 1 \ . \tag{2}$$

Note that the problem is symmetric in receivers and transmitters, i.e. if each receiver instead could be viewed as having a common direction to all transmitters, the same problem can be solved for transmitter positions and receiver directions. We denote $\mathbf{f} = [f_1,\dots,f_m]^T$, $\mathbf{g} = [g_1,\dots,g_k]$, $\mathbf{r} = [\mathbf{r}_1,\dots,\mathbf{r}_m]$ and $\mathbf{n} = [\mathbf{n}_1,\dots,\mathbf{n}_k]$.

The problem of determining full transmitter positions $\mathbf{b}_j$ instead of directions $\mathbf{n}_j$, see (1), seems harder than using the far field approximation as in Problem 2.1. The measurements are now bilinearly dependent on $\mathbf{r}_i$ and $\mathbf{n}_j$. Algorithms that explicitly consider the far field assumption are also required, as the problem of determining general positions of transmitters when the far field approximation is in effect, is an ill-conditioned problem.

We denote the problem as minimal if the number of solutions for generic distance measurements δ_{ij} is finite and positive, disregarding solutions that are the same up to gauge freedom.

2.2 Gauge Freedom

The unknown parameters $(\mathbf{r}, \mathbf{n}, \mathbf{f}, \mathbf{g})$ have certain degrees of freedom that does not change the measurements, called gauge freedom. Any translation $\mathbf{t}$, rotation matrix $\mathbf{R}$ and offset change K can be applied to the solution according to

$$\begin{aligned} \mathbf{r}_{i,\text{trans}} &= \mathbf{r}_i + \mathbf{t}, \ g_{j,\text{trans}} = g_j - \mathbf{t}^T \mathbf{n}_j \\ \mathbf{r}_{i,\text{rot}} &= \mathbf{R}\mathbf{r}_i, \ \mathbf{n}_{j,\text{rot}} = \mathbf{R}\mathbf{n}_j \\ f_{i,\text{offs}} &= f_i + K, \ g_{j,\text{offs}} = g_j - K \end{aligned}$$

without changing the measurements δ_{ij}. Thus, we can only hope to reconstruct the unknowns up to these seven degrees of freedom.

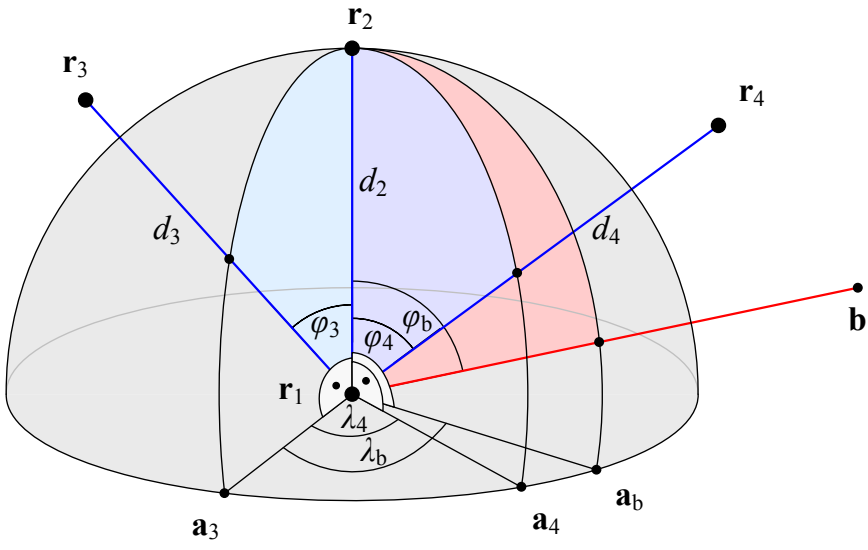


Figure 1: Scheme of the Ellipsoid method. Three distances d_2, d_3, d_4, and three angles φ_3, φ_4, and ϑ define a tetrahedron of four receivers $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, $\mathbf{r}_4$. Transmitter $\mathbf{b}$ is assumed to be far away from the receivers. Its signal arrives from the angles φ_b, λ_b.

3 The Ellipsoid Method in Three-Dimensional Space

We propose the Ellipsoid TDOA method which solves the FFUTOA calibration problem for four receivers using at least nine transmitters. The time differences of signals from distant emitters form an ellipsoid which characterizes the distances and angles between four receivers. An elegant representation can be derived from the knowledge that an ellipsoid corresponds to a covariance matrix. Once this covariance matrix is known, one can extract the parameters that generate the ellipsoid from the matrix, i.e. the configuration of four receivers.

3.1 Definition of the Covariance Ellipsoid

A rigid tetrahedron of four receivers is defined by three distances $d_2 = \|\mathbf{r}_1 - \mathbf{r}_2\|$, $d_3 = \|\mathbf{r}_1 - \mathbf{r}_3\|$, $d_4 = \|\mathbf{r}_1 - \mathbf{r}_4\|$, two height angles $\varphi_3 = \angle_{\mathbf{r}_2\mathbf{r}_1\mathbf{r}_3}$, $\varphi_4 = \angle_{\mathbf{r}_2\mathbf{r}_1\mathbf{r}_4}$, and the azimuth angle $\lambda_4 = \angle_{\mathbf{a}_3\mathbf{r}_1\mathbf{a}_4}$, see Fig. 1. Furthermore we define $\vartheta = \angle_{\mathbf{r}_3\mathbf{r}_1\mathbf{r}_4}$.

A signal arrives from the angles $\varphi_b = \angle_{\mathbf{r}_2\mathbf{r}_1\mathbf{b}}$ and $\lambda_b = \angle_{\mathbf{a}_3\mathbf{r}_1\mathbf{a}_b}$, uniquely determining the direction. The signal angles with respect to two receivers are $\gamma_2 = \angle_{\mathbf{r}_2\mathbf{r}_1\mathbf{b}}$, $\gamma_3 = \angle_{\mathbf{r}_3\mathbf{r}_1\mathbf{b}}$, and $\gamma_4 = \angle_{\mathbf{r}_4\mathbf{r}_1\mathbf{b}}$. Omitting the signal index, these angles are defined by the UTOA measures according to the cosine law as

$$\begin{aligned} x = \delta_1 - \delta_2 = d_2\cos(\gamma_2)\ , \quad y = \delta_1 - \delta_3 = d_3\cos(\gamma_3)\ , \\ \text{and} \quad z = \delta_1 - \delta_4 = d_4\cos(\gamma_4)\ . \end{aligned} \tag{3}$$

The auxiliary points $\mathbf{a}_3$, $\mathbf{a}_4$, and $\mathbf{a}_b$ are projections of $\mathbf{r}_3$, $\mathbf{r}_4$, and $\mathbf{b}$ respectively, onto the plane orthogonal to $\mathbf{r}_1 - \mathbf{r}_2$ through $\mathbf{r}_1$.

In the following we derive the covariance matrix for time differences in the Eqns. (3) assuming uniform signal source positions. This matrix characterizes a covariance ellipsoid, [21], describing the ellipsoid which the time differences reside on. If this matrix is known, the distances and angles between the receivers can be directly read from the matrix. We state the following definition.

Definition 3.1. The $\boldsymbol{\Sigma}$-ellipsoid for covariance matrix $\boldsymbol{\Sigma}$ is the ellipsoid with center $\boldsymbol{\mu}$ where for all points $\mathbf{x}$ holds

$$d_{\text{Mah}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma}) \ = \ \sqrt{(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})} \ = \ 1\ .$$

The metric $d_{\text{Mah}}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the Mahalanobis distance. For $\boldsymbol{\Sigma}$-ellipsoids the following holds.

Lemma 3.1. *The covariance of points uniformly distributed over a $\boldsymbol{\Sigma}$-ellipsoid in $\mathbb{R}^3$ is $\hat{\boldsymbol{\Sigma}} = \frac{1}{3}\boldsymbol{\Sigma}$. In the two-dimensional case the covariance is $\hat{\boldsymbol{\Sigma}} = \frac{1}{2}\boldsymbol{\Sigma}$.*

Lemma 3.1 can be verified by integration over all points of the $\boldsymbol{\Sigma}$-ellipsoid and calculating the covariance. Given the definition of the covariance ellipsoid we propose the following theorem.

Theorem 3.1. *The time differences (x, y, z) of distant signals arriving at four receivers $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, $\mathbf{r}_4$ in space $\mathbb{R}^3$ form a $3\hat{\boldsymbol{\Sigma}}$-ellipsoid with covariance matrix*

$$\hat{\boldsymbol{\Sigma}} \ = \ \frac{1}{3}\begin{pmatrix} d_2^2 & d_2 d_3\cos(\varphi_3) & d_2 d_4\cos(\varphi_4) \\ d_2 d_3\cos(\varphi_3) & d_3^2 & d_3 d_4\cos(\vartheta) \\ d_2 d_4\cos(\varphi_4) & d_3 d_4\cos(\vartheta) & d_4^2 \end{pmatrix}\ .$$

Proof. The proof is directly based on the definition of a covariance ellipsoid. The first thing to show is that the matrix $\hat{\mathbf{\Sigma}}$ is actually a covariance matrix, therefore is positive semi-definite. For simplicity we assume that the receivers are synchronized, therefore the mean $\boldsymbol{\mu}$ is zero. In case they are not, synchronize the receivers by regression as described in the next Section 3.2.

Now, consider the continuous distribution of synchronized time differences over uniformly distributed directions of origin. Such a uniform distribution of signal origins $\hat{\mathbf{b}}$ in space $\mathbb{R}^3$ can be created by points

$$\hat{\mathbf{b}} \;=\; R \cdot \big(r\cos(\lambda),\; r\sin(\lambda),\; \ell\big)^T \;,$$

where $\lambda \in [0, 2\pi]$ and $\ell \in [-1, 1]$ are uniformly independently distributed random variables, and $r = \sqrt{1-\ell^2}$. The density function of the distribution is $h(\lambda, \ell) = g(\lambda)f(\ell) = \frac{1}{4\pi}$. Without loss of generality, the tetrahedron is aligned such that $\mathbf{r}_1$ is the origin, $\mathbf{r}_2$ is parallel to the $\hat{z}$-axis, and $\mathbf{r}_3$ resides on the $\hat{x}/\hat{z}$-plane. Assuming that the sphere is large, i.e. the signals $\hat{\mathbf{b}}$ originate from far away, the angles of the signals are

$$\lambda_\mathrm{b} = \lambda \qquad \text{and} \qquad \cos(\varphi_\mathrm{b}) = \ell \;. \tag{4}$$

By using spherical trigonometry and the Eqns. (3) we calculate the time differences $\hat{\mathbf{x}} = [x, y, z]^T$ with respect to the tetrahedron angles as follows

$$\begin{aligned} x \;&=\; d_2\cos(\gamma_2) \;&=\; d_2\Big(\cos(\varphi_\mathrm{b})\Big) \\ y \;&=\; d_3\cos(\gamma_3) \;&=\; d_3\Big(\cos(\varphi_3)\cos(\varphi_\mathrm{b}) + \sin(\varphi_3)\sin(\varphi_\mathrm{b})\cos(\lambda_\mathrm{b})\Big) \\ z \;&=\; d_4\cos(\gamma_4) \;&=\; d_4\Big(\cos(\varphi_4)\cos(\varphi_\mathrm{b}) + \sin(\varphi_4)\sin(\varphi_\mathrm{b})\cos(\lambda_\mathrm{b} - \lambda_4)\Big) \;. \end{aligned} \tag{5}$$

Note that the angles γ_2, γ_3, γ_4, are not uniformly distributed in the three-dimensional case, in contrast to the planar case. We express ϑ as

$$\cos(\vartheta) \;=\; \sin(\varphi_3)\sin(\varphi_4)\cos(\lambda_4) + \cos(\varphi_3)\cos(\varphi_4) \;. \tag{6}$$

Using the uniform distribution of signals (4) and the time differences $\hat{\mathbf{x}}$ from Eqns. (5) that follow, we show by integration that the time differences characterize a covariance matrix as stated in Theorem 3.1:

$$\hat{\mathbf{\Sigma}} \quad = \quad \int_0^{2\pi}\!\!\int_{-1}^{1} \hat{\mathbf{x}}\,\hat{\mathbf{x}}^T\, h(\gamma,\ell)\;\mathrm{d}\ell\;\mathrm{d}\lambda \;=\; h(\gamma,\ell)\int_0^{2\pi}\!\!\int_{-1}^{1} \begin{pmatrix} x^2 & xy & xz \\ xy & y^2 & yz \\ xz & yz & z^2 \end{pmatrix}\;\mathrm{d}\ell\;\mathrm{d}\lambda$$

$$\overset{(4)-(6)}{=} \quad \frac{1}{3}\begin{pmatrix} d_2^2 & d_2 d_3\cos(\varphi_3) & d_2 d_4\cos(\varphi_4) \\ d_2 d_3\cos(\varphi_3) & d_3^2 & d_3 d_4\cos(\vartheta) \\ d_2 d_4\cos(\varphi_4) & d_3 d_4\cos(\vartheta) & d_4^2 \end{pmatrix} . \tag{7}$$

Due to the quadratic form is $\hat{\mathbf{\Sigma}}$ positive semidefinite. Furthermore, the matrix is definite, which follows from the fact that the time differences are bounded.

The next step is to verify that the time differences are actually characterized by the matrix. The distribution of signal directions $(\lambda_\mathrm{b}, \varphi_\mathrm{b})$ is irrelevant for this step, and for application of the algorithm. However, as the points $\hat{\mathbf{b}}$ in Eq. (4) cover the complete sphere, all signal directions are considered. Calculating the Mahalanobis distance by inserting $\hat{\mathbf{x}}$ and $\hat{\mathbf{\Sigma}}$ yields

$$d_\mathrm{Mah}\left(\hat{\mathbf{x}}, \vec{0}, 3\hat{\mathbf{\Sigma}}\right) \;=\; \sqrt{\hat{\mathbf{x}}^T\left(3\hat{\mathbf{\Sigma}}\right)^{-1}\hat{\mathbf{x}}} \qquad\qquad = \; 1 \; ,$$

revealing that all time difference points have constant Mahalanobis distance from the origin, therefore reside on an ellipsoid, which is according to Lemma 3.1 the $3\hat{\mathbf{\Sigma}}$-ellipsoid. □

3.2 Transformation of the Covariance Matrix

We now describe the transformation of parameters from a regression polynomial to the parameters of the covariance matrix. Under the assumption of a zero-mean ellipsoid, i.e. the receivers are synchronized, an ellipsoid is described by a polynomial equation

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz = 1 \; . \tag{8}$$

Regression of at least $m \geq 6$ signals in the equation system

$$\underbrace{\begin{pmatrix} x_1^2 & y_1^2 & z_1^2 & x_1y_1 & x_1z_1 & y_1z_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_m^2 & y_m^2 & z_m^2 & x_my_m & x_mz_m & y_mz_m \end{pmatrix}}_{\mathbf{Q}} \underbrace{(a,b,c,d,e,f)^T}_{\mathbf{u}} \;=\; \vec{1}$$

and solving a least squares scheme for $\mathbf{u} \;=\; \left(\mathbf{Q}^T\mathbf{Q}\right)^{-1}\left(\mathbf{Q}^T\vec{1}\right)$ yields ellipsoid parameters a to f.

An ellipsoid in space $\mathbb{R}^3$ can be represented by the matrix form $\mathbf{x}^T\mathbf{\Sigma}^{-1}\mathbf{x} = 1$, where $\mathbf{x} = [x, y, z]^T$ is a vector and $\mathbf{\Sigma}$ is a symmetric positive definite matrix

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_1^2 & \omega_1 & \omega_2 \\ \omega_1 & \sigma_2^2 & \omega_3 \\ \omega_2 & \omega_3 & \sigma_3^2 \end{pmatrix} .$$

Substitution and conversion of the parameter set yields the parameters of the covariance matrix

$$\begin{aligned} \sigma_1^2 &= (f^2 - 4bc) \,/\, Z & \omega_1 &= (2cd - ef) \,/\, Z \\ \sigma_2^2 &= (e^2 - 4ac) \,/\, Z & \omega_2 &= (2be - df) \,/\, Z \\ \sigma_3^2 &= (d^2 - 4ab) \,/\, Z & \omega_3 &= (2af - de) \,/\, Z \end{aligned} \tag{9}$$

where $Z = be^2 + cd^2 + af^2 - 4abc - def$.

In case the receivers are not synchronized, the ellipsoid is shifted to zero-mean by converting the general ellipsoid polynomial equation to a translation-invariant form. In three dimensions the general form is

$$ax^2 + by^2 + cz^2 + dxy + exz + fyz + gx + hy + jz = 1 \,, \tag{10}$$

for which the parameters a to j are calculated by regression of at least nine signals. The parameters are converted to the following translation-invariant form

$$\begin{aligned} &\hat{a}(x - \hat{u})^2 + \hat{b}(y - \hat{v})^2 + \hat{c}(z - \hat{w})^2 \\ &+ \hat{d}(x - \hat{u})(y - \hat{v}) + \hat{e}(x - \hat{u})(z - \hat{w}) + \hat{f}(y - \hat{v})(z - \hat{w}) = 1. \end{aligned} \tag{11}$$

Calculation of $\hat{a}$ to $\hat{f}$ and $\hat{u}$, $\hat{v}$, $\hat{w}$ from the coefficients of Eq. (10) can be done in a computer algebra software by expansion of Eq. (11) and substitution of the constant term. The coefficients $\hat{a}$ to $\hat{f}$ are converted for the covariance matrix using Eqns. (9). The coefficient vector $(\hat{u}, \hat{v}, \hat{w})^T$ equals the center point of the ellipse and the synchronization offset of the receivers.

According to Theorem 3.1, the distances and angles in the tetrahedron of receivers are now directly characterized by the coefficients of the covariance matrix. The distances and angles are calculated by

$$\begin{aligned} d_2 &= \sqrt{3}\,\sigma_1 \,, & \cos(\varphi_3) &= \tfrac{\omega_1}{\sigma_1\sigma_2} \,, \\ d_3 &= \sqrt{3}\,\sigma_2 \,, & \cos(\varphi_4) &= \tfrac{\omega_2}{\sigma_1\sigma_3} \,, \\ d_4 &= \sqrt{3}\,\sigma_3 \,, & \cos(\vartheta) &= \tfrac{\omega_3}{\sigma_2\sigma_3} \,. \end{aligned}$$

3.3 Degenerate Cases

When measurements δ_{ij} are corrupted by noise, or the far field assumption is violated, the solution of parameters in (10) might not yield an ellipsoid, but another type of quadric surface. For four receivers and nine transmitters, this constitutes a case of given measurements where there is no exact real solution to (2), as such time differences must lie on an ellipsoid by Theorem 3.1. Instead of using the regression scheme, one can obtain an approximation based on Theorem 3.1 by covariance estimation of the given time differences (x, y, z), denoted $\mathbf{\Sigma}^*$. Using $\hat{\mathbf{\Sigma}} = \frac{1}{3}\mathbf{\Sigma}^*$, distance and angle parameters can be estimated as in Section 3.2.

Other degenerate cases are when the ellipsoid is collapsed to a ellipse surface, or when transformed time differences in (11) lie on two intersecting quadric surfaces, thus giving infinite number of solutions.

4 Matrix Factorization Method

The matrix factorization method uses linear techniques to solve Problem 2.1 for receiver positions, transmitter directions and offsets. At least four receivers and nine transmitters are needed. Without loss of generality we assume that the solution is partially normalized for gauge freedom as the first receiver $\mathbf{r}_1 = \mathbf{0}$ and $f_1 = 0$, see Section 2.2.

Using the FFUTOA measurements δ_{ij}, collected in the matrix $\tilde{\mathbf{D}} = [\delta_{ij}]_{m\times k}$ we immediately obtain the unknowns g_j since $\delta_{1j} = \mathbf{r}_1^T\mathbf{n}_j + f_1 + g_j = g_j$, since $\mathbf{r}_1 = \mathbf{0}$ and $f_1 = 0$. We then subtract the first row containing g_j from all other rows of $\tilde{\mathbf{D}}$ and remove the first row of zeros to obtain a new matrix that fulfill

$$\mathbf{D}_2 = \begin{bmatrix}\mathbf{r}^T & \mathbf{f}\end{bmatrix}\begin{bmatrix}\mathbf{n} \\ \vec{1}\end{bmatrix} \tag{12}$$

where $\vec{1}$ is a vector of ones. $\mathbf{D}_2$ is a product of two matrices of rank ≤ 4 and is thus itself of rank ≤ 4. This is used in [18]. Here we further reduce the rank of the factorization by subtracting the first column of $\mathbf{D}_2$ from all the other columns and remove the first row of columns. Both steps manipulating $\tilde{\mathbf{D}}$ can be done

using the compaction matrices $\mathbf{C}_m$ of size $(m-1) \times m$ and $\mathbf{C}_k$ of size $k \times (k-1)$:

$$\mathbf{C}_m = \begin{bmatrix} -1 & 1 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \dots & 1 \end{bmatrix}, \; \mathbf{C}_k = \begin{bmatrix} -1 & -1 & \dots & -1 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}. \tag{13}$$

Then we have

$$\mathbf{D} = \mathbf{C}_m \tilde{\mathbf{D}} \mathbf{C}_k = \begin{bmatrix} \tilde{\mathbf{r}}^T & \mathbf{f} \end{bmatrix} \begin{bmatrix} \tilde{\mathbf{n}} \\ \vec{0} \end{bmatrix} = \tilde{\mathbf{r}}^T \tilde{\mathbf{n}}, \tag{14}$$

where $\tilde{\mathbf{r}}$ equals $\mathbf{r}$ with the first receiver removed as $\mathbf{r}_1 = \mathbf{0}$, and $\tilde{\mathbf{n}}$ is a $3 \times (k-1)$ matrix with the j^{th} column $\tilde{\mathbf{n}}_j = \mathbf{n}_{j+1} - \mathbf{n}_1$. Now we have a rank-3 factorization, thus requiring at least four receivers and four transmitters. After applying SVD to $\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ we obtain the rank-3 factorization such that $\mathbf{D} = \bar{\mathbf{r}}^T \bar{\mathbf{n}}$ where $\bar{\mathbf{r}} = \mathbf{U}_3 \mathbf{S}_3$ and $\bar{\mathbf{n}} = \mathbf{V}_3^T$. $\mathbf{U}_3$, $\mathbf{S}_3$ and $\mathbf{V}_3$ are the truncated parts of the SVD corresponding to the three largest singular values. This factorization of $\mathbf{D}$ is unique up to an unknown transformation $\mathbf{H}$ i.e. $\mathbf{D} = \bar{\mathbf{r}}^T \mathbf{H}^{-1} \mathbf{H} \bar{\mathbf{n}}$. We will find $\tilde{\mathbf{n}}_j = \mathbf{H}\bar{\mathbf{n}}$ i.e. $\mathbf{n}_{j+1} - \mathbf{n}_1 = \mathbf{H}\bar{\mathbf{n}}_j$ by using the constraints that

$$\begin{aligned} \tilde{\mathbf{n}}_j^T \tilde{\mathbf{n}}_j &= (\mathbf{n}_{j+1} - \mathbf{n}_1)^T (\mathbf{n}_{j+1} - \mathbf{n}_1) = 2 - 2\mathbf{n}_{j+1}^T \mathbf{n}_1 \\ &= 2 - 2(\mathbf{H}\bar{\mathbf{n}}_j + \mathbf{n}_1)^T \mathbf{n}_1 = -2\bar{\mathbf{n}}_j^T \mathbf{H}^T \mathbf{n}_1 = \bar{\mathbf{n}}_j^T \mathbf{H}^T \mathbf{H} \bar{\mathbf{n}}_j \, . \end{aligned} \tag{15}$$

We apply a change of variables with a 3×3 symmetric $\mathbf{C} = \mathbf{H}^T\mathbf{H}$ and a 3×1 vector $\mathbf{v} = \mathbf{H}^T \mathbf{n}_1$. From (15), we have the following equation for transmitter j:

$$\bar{\mathbf{n}}_j^T \mathbf{C} \bar{\mathbf{n}}_j + 2\bar{\mathbf{n}}_j^T \mathbf{v} = 0. \tag{16}$$

These equations are linear in the elements of $\mathbf{C}$ and $\mathbf{v}$ which have in total 9 variables. In general, with 8 such equations (thus 9 transmitters), we can solve this homogeneous linear equation system uniquely up to scale.

We can extract the solutions for $\mathbf{C}$ and $\mathbf{v}$ from the solution to the linear equation which is valid up to an unknown scaling factor and sign. We can determine the sign by using that $\mathbf{C}$ is positive definite and compute $\mathbf{H}$ by applying Cholesky factorization $\mathbf{C} = \mathbf{H}^T\mathbf{H}$. As $\mathbf{H}^T\mathbf{H} = \mathbf{H}^T\mathbf{R}^T\mathbf{R}\mathbf{H}$ for a rotation/mirroring matrix $\mathbf{R}$, this will give $\mathbf{H}$ uniquely up to $\mathbf{R}$. But as $\mathbf{R}$ corresponds to rotating/mirroring

Algorithm 1 Minimal solver for the FFUTOA calibration problem

Input: FFUTOA measurement matrix $\tilde{\mathbf{D}}$ of size $(m = 4) \times (k = 9)$.

Output: Receiver positions $\mathbf{r}$, transmitter directions $\mathbf{n}$, receiver offsets $\mathbf{f}$ and transmitter offsets $\mathbf{g}$.

Postconditions: (i) $\mathbf{D}$ must have rank 3, (ii) the linear equations (16) must only have a null space of dimension one, (iii) $\mathbf{C}$ must be positive definite.

1: Set $g_j := \tilde{\mathbf{D}}_{1j}$ and $\mathbf{D} := \mathbf{C}_m \tilde{\mathbf{D}} \mathbf{C}_k$ where $\mathbf{C}_l, \mathbf{C}_m$ is the compaction matrices in (13)
2: Calculate the SVD $\mathbf{D} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ and set $\bar{\mathbf{r}}$ to first three columns of $\mathbf{US}$ and $\bar{\mathbf{n}}$ to first three rows of $\mathbf{V}^T$
3: For the unknowns in the symmetric matrix $\mathbf{C}$ and vector $\mathbf{v}$, get the solution space for the equations $\bar{\mathbf{n}}_j^T \mathbf{C} \bar{\mathbf{n}}_j + 2\bar{\mathbf{n}}_j^T \mathbf{v} = 0$ where $\bar{\mathbf{n}}_j$ is the j^{th} column of $\bar{\mathbf{n}}$
4: Set the sign of the solution $\mathbf{C}, \mathbf{v}$ such that $C_{11} > 0$
5: Calculate the Cholesky decomposition $\mathbf{C} = \mathbf{H}^T\mathbf{H}$
6: Lock the scale of the solutions $\mathbf{H}, \mathbf{v}$ so that $\|\mathbf{H}^{-T}\mathbf{v}\| = 1$
7: Set $\mathbf{n}_1 := \mathbf{H}^{-T}\mathbf{v}$, $\mathbf{n}_{j+1} := \mathbf{H}\bar{\mathbf{n}}_j + \mathbf{n}_1$ and $\mathbf{r} := \mathbf{H}^{-T}\hat{\mathbf{r}}$

the coordinate system, $\mathbf{R}$ is a gauge freedom according to Section 2.2 and can be set to the identity matrix.

We can find the scale by using the constraint $\|\mathbf{n}_1\| = \|\mathbf{H}^{-T}\mathbf{v}\| = 1$. Note that fixing the scale in this way will also guarantee that $\mathbf{n}_j^T\mathbf{n}_j = (\mathbf{H}^T\bar{\mathbf{n}}_j + \mathbf{n}_1)^T(\mathbf{H}^T\bar{\mathbf{n}}_j + \mathbf{n}_1) = \underbrace{\bar{\mathbf{n}}_j^T\mathbf{H}^T\mathbf{H}\bar{\mathbf{n}}_j + 2\bar{\mathbf{n}}_j^T\mathbf{H}^T\mathbf{n}_1}_{=0 \text{ by (15)}} + \mathbf{n}_1^T\mathbf{n}_1 = \mathbf{n}_1^T\mathbf{n}_1 = 1$. Summarizing these steps yields Algorithm 1.

4.1 Degenerate Cases

Theorem 4.1. *Degenerate cases for the minimal algorithm are when i) The transformed measurement matrix* $\mathbf{D}$ *has Rank*$(\mathbf{D}) \leq 2$ *or ii) The difference of the transmitter directions* $\mathbf{n}_j - \mathbf{n}_1$ *lie on the intersection of two or more quadric surfaces with constant term 0.*

Case i) happens iff receivers or transmitter directions lie in a plane. For ii), the case when the transmitter directions $\mathbf{n}_j$ *lie on the intersection of two or more a quadric surfaces is a special case.*

Proof. The only time the algorithm fails is when the prerequisites are not fulfilled.

This happens iff i) $Rank(\mathbf{D}) \leq 2$ or ii) the linear equations (16) have a null space of dimension two or more.

For case i), step 2 will extract data from the SVD that are not uniquely determined from the measurements, but has several degrees of freedom. This will result in a reconstruction of $\mathbf{r}, \mathbf{n}$ that fulfills the measurements, but is not unique, as there are an infinite number of solutions.

$Rank(\mathbf{D}) \leq 2$ iff either receivers $\mathbf{r}$ or difference of transmitter directions $\mathbf{n}_j - \mathbf{n}_1$ in (14) are embedded in a lower dimensional subspace than assumed. Remembering that receiver positions can be translated as in Section 2.2, this is equivalent to receivers or transmitter directions being embedded in a plane.

For case (ii), there are at least two non linearly dependent solutions to (16). The solutions can be seen as constants for a quadric surface with radius 0 that $\bar{\mathbf{n}}_j$ should lie on, i.e.

$$\bar{\mathbf{n}}_j^T \mathbf{C}_1 \bar{\mathbf{n}}_j + \bar{\mathbf{n}}_j^T \mathbf{D}_1 = 0, \ \bar{\mathbf{n}}_j^T \mathbf{C}_2 \bar{\mathbf{n}}_j + \bar{\mathbf{n}}_j^T \mathbf{D}_2 = 0,$$

where $[\mathbf{C}_1 \ \mathbf{D}_1] \neq \lambda [\mathbf{C}_2 \ \mathbf{D}_2]$ for all $\lambda \in \mathbb{R} \backslash \{0\}$ and $\mathbf{C}_i$ symmetric. As $\mathbf{n}_{j+1} - \mathbf{n}_1 = \mathbf{H}\bar{\mathbf{n}}_j$, this is equivalent to

$$\begin{aligned} \left(\mathbf{n}_j - \mathbf{n}_1\right)^T \mathbf{H}^{-T} \mathbf{C}_1 \mathbf{H}^{-1} \left(\mathbf{n}_j - \mathbf{n}_1\right) + \left(\mathbf{n}_j - \mathbf{n}_1\right)^T \mathbf{H}^{-T} \mathbf{D}_1 = 0, \\ \left(\mathbf{n}_j - \mathbf{n}_1\right)^T \mathbf{H}^{-T} \mathbf{C}_2 \mathbf{H}^{-1} \left(\mathbf{n}_j - \mathbf{n}_1\right) + \left(\mathbf{n}_j - \mathbf{n}_1\right)^T \mathbf{H}^{-T} \mathbf{D}_2 = 0, \end{aligned} \tag{17}$$

which is equivalent of the difference of the receiver directions $\mathbf{n}_j - \mathbf{n}_1$ lying on two or more quadric surfaces with constant term 0. As a special case, if the transmitter directions $\mathbf{n}_j$ lie on two or more different quadric surfaces, then the differences $\mathbf{n}_j - \mathbf{n}_1$ will fulfill (17). □

Note that the degenerate cases characterized in i) is inherent to the problem, not the algorithm. There are fewer degrees of freedom to estimate than assumed, and thus there is not a unique solution. If both receivers and transmitter directions lie in the same plane, a similar algorithm for 2D based on the same factorization steps and equations can readily be constructed.

A special case is when $\mathbf{C}$ is not positive definite. Then there exists no real factorization $\mathbf{C} = \mathbf{H}^T\mathbf{H}$. There exists complex factorizations, e.g. obtained using eigenvalue decomposition $\mathbf{C} = \mathbf{Q}^T\mathbf{D}\mathbf{Q} = \mathbf{Q}^T\sqrt{\mathbf{D}}^T\sqrt{\mathbf{D}}\mathbf{Q} = \mathbf{H}^T\mathbf{H}$, which results in complex solutions. These cases equate exactly to the cases where the ellipsoid method does not get an ellipsoid from solving (10), as these are the cases where there are no exact real solutions to the given measurements.

5 Extension to Overdetermined Cases and Noise

Both algorithms solve a minimal case, meaning that there are only a finite positive number of solutions to (2) given arbitrary measurements in general enough position. This can be seen from the fact that the matrix factorization algorithm does not lose any solutions from the solution space by any particular choice in any of the steps. Thus there one solution discounting gauge freedom. Another way of seeing it is by counting degrees of freedom. When using $m = 4$ receivers and $k = 9$ transmitters, the number of measurements $mk = 36$ equals to the number of unknowns, $4m + 3k - 7 = 36$ accounting for gauge freedom.

When having more than four receivers, more than nine transmitters and the measurements d_{ij} are not true FFUTOA measurements, due to noise or that the far field assumption does not hold, both methods can be extended in a straightforward manner.

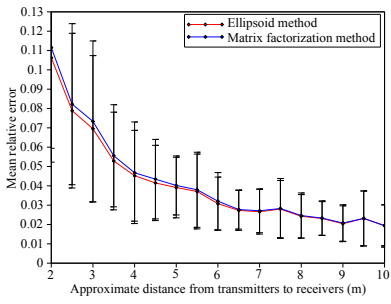
For the ellipsoid method, two modifications are made. (i) When having more than nine receivers, the least squares solution to (11) can be calculated. (ii) When having more than four receivers, subproblems using only four receivers at a time are solved. With overlap of receivers used in the different subproblems, all distances between receivers can be calculated and multidimensional scaling [22] can be used to get the full coordinates of all receivers.

For the matrix factorization method, the three following modifications are made. (i) In step 2, the best rank 3 approximation can still be obtained by SVD, although $\mathbf{D}$ is not necessarily rank 3. (ii) The system of equations in step 3 will in general only have the trivial solution, but is approximated to rank 8 by SVD to still attain the expected one dimensional solution set. (iii) $\|\mathbf{n}_j\|$ is only approximately 1, so $\mathbf{n}_j$ is normalized to be of length 1.

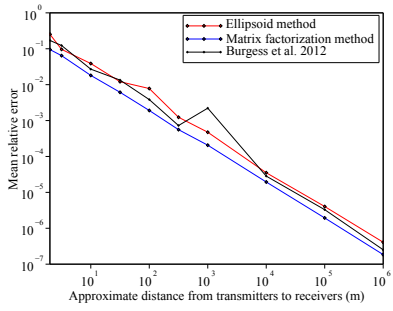
From here on, the extended methods will be used. Note that when only minimal number of measurements are available, the extended methods are equivalent to the minimal ones.

6 Experimental Validation

To be able to evaluate the quality of a solution, receivers positions $\mathbf{r}_i$ are compared to ground truth receiver positions $\mathbf{r}_{i,gt}$. Receivers are rotated, mirrored and translated so that $\sum_i \|\mathbf{R}(\mathbf{r}_i - \mathbf{t}) - \mathbf{r}_{i,gt}\|^2$ is minimized, where $\mathbf{R}$ and $\mathbf{t}$ is a rotation/mirroring and translation respectively. Finding $\mathbf{R}$ and $\mathbf{t}$ is done by using [23].



(a) Far field test, 4 receivers, 15 transmitters.



(b) Far field test, 5 receivers, 15 transmitters.

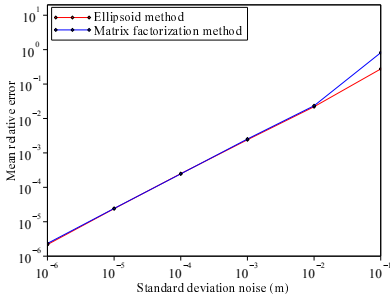
Figure 2: Mean relative error of reconstructed receiver positions for 100 runs, plotted against the approximate distance from receivers to transmitters. (a) Bars are ± 1 standard deviation for the different transmitter distances.

For all experiments, relative errors are then defined as $\|\mathbf{r} - \mathbf{r}_{gt}\|_{Fro} / \|\mathbf{r}_{gt}\|_{Fro}$ where $\|\cdot\|_{Fro}$ is the Frobenius norm. All algorithms were implemented and run on a standard desktop computer in Scilab.

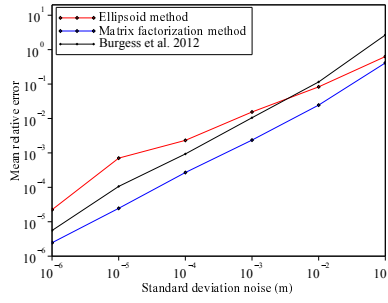
6.1 Simulations

For all simulations, offsets f_i and g_j are drawn from i.i.d. uniform distributions over $[0, 10]$. To evaluate the assumption that transmitters have a common direction to the receivers, transmitter positions $\mathbf{b}_j$ were uniformly distributed over a sphere of radius d. To be able to control how much further away transmitters were from receivers than the inter distance between receivers, four receivers were placed at a tetrahedron around the origin with side length 1m. As signal sources are often easily obtained in applications, 15 transmitters were used. UTOA Measurements were constructed as $\delta_{ij} = \|\mathbf{r}_i - \mathbf{b}_j\| + f_i + g_j$. The mean relative error for 100 runs each plotted against the transmitter distance d to the origin can be seen for different radii d in Fig. 2a for the minimal four receivers and in Fig. 2b for five receivers. The extra receiver was uniformly distributed in the cube of which the tetrahedron of the four first receivers were inscribed to.

Fig. 2a shows that using only four receivers, both algorithms can get under 5% relative error with having transmitter approximately four times further away



(a) 4 receivers, 15 transmitters.



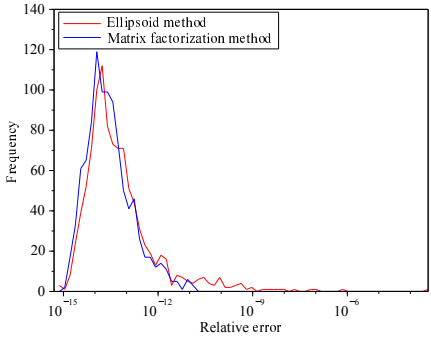
(b) 5 receivers, 15 transmitters.

Figure 3: Measurements with additive Gaussian white noise. The standard deviation is plotted against the mean relative error of reconstructed receiver positions for 100 runs.

than the inter distance between receivers. For the experiment in Fig. 2b, we compare the results to the method in [18] as we now have the five receivers for the method to be applicable. The results indicate the ellipsoid method being slightly worse on short distances and the matrix factorization method being generally more accurate. Mean execution time was 8.0 ms, 2.1 ms and 30 ms for the ellipsoid method, matrix factorization method and the method in [18] each.

To test the robustness of the methods, white Gaussian noise was added to the measurements. The same setup as for the far field experiments was used, with transmitter distance of 10^7 from the receivers. In Fig. 3 relative error of reconstructed receiver positions are plotted against the standard deviation of the noise. The results indicate the ellipsoid method being slightly better with higher noise level when using the minimum four receivers, and the matrix factorization outperforming both the ellipsoid and the method in [18] using five receivers.

The numerical performance of the minimal algorithms were evaluated by generating problems where the measurement matrix $\tilde{\mathbf{D}}$ are FFUTOA (2). Receivers are drawn from i.i.d. uniform distributions in a cube of unit volume centered around the origin. Nine transmitter directions and four receivers were simulated. The error distribution for 1000 such experiments can be seen in Fig. 4a. Mean execution time for the ellipsoid method and the matrix factorization method was 3.2 and 1.9 ms respectively.



(a) Minimal solver numerical performance.

(b) Experimental setup

Figure 4: (a) Numerical performance of minimal solver in 1000 simulated experiments. (b) Setup for indoor experiment using microphones and distinct manually made sounds.

6.2 Real Data

The same data as in [18] was used, where the measurements d_{ij} were obtained from an experimental setup using eight SHURE SV100 microphones as receivers and random distinct manually made sounds as transmitters. The microphones were connected to a M-Audio Fast Track Ultra 8R audio interface. The 19 sound sources were approximately 30 m away from the receivers. Microphones were set in the corners of a cuboid of roughly $100 \times 105 \times 60\,\text{cm}^3$. A picture of the experiment setup can be seen in Fig. 4b. The microphone offsets were created by adding uniformly i.i.d. silences between 0-1 s long to the beginning of each sound track, effectively starting the recordings at different unknown times. The beginning of each sound were matched by a heuristic cross correlation algorithm to create TDOA measurements.

As we have more than five microphones, the algorithms were also compared using the method in [18]. The mean reconstruction error on the microphone positions were 15 cm, 5 cm and 14 cm for the ellipsoid method, matrix factorization method and the method in [18] respectively. Most of the error are in the floor-to-roof direction. This can be explained by the sounds all being made close to ground level and thus the transmitter directions will be close to being in a plane, giving poor resolution in floor-to-roof direction.

7 Conclusions

We have presented two methods for solving the previously unsolved problem of sensor network calibration using only a minimal number of unsynchronized TDOA measurements in a far field setting. The assumption of far field signals is important, as the problem of trying to determine exact positions for transmitters is ill conditioned when the far field assumption is close to true.

Simulated experiments support the feasibility of the methods, and show that the minimal algorithms are numerically stable and fast, making them suitable in RANSAC schemes to weed out outliers. They also handle additive Gaussian noise well. The far field assumption gives good results as long as transmitter-receiver distances are four times larger than inter-receiver distances.

A comparison between the two methods, running on the minimal case of four receivers and nine transmitters, indicates the matrix factorization method being slightly faster and having better worst case precision than the ellipsoid method. The ellipsoid method however has a more plausible way of handling the case when the measurements are such that no exact real solutions exist, as per Section 3.3 and 4.1. The ellipsoid method estimates the covariance of the time differences for parameter estimation, whereas the matrix factorization finds a complex solution.

When having more than the minimum amount of four receivers and nine transmitters, the matrix factorization is easily extended to handle more than the minimal number of receivers and transmitters, and usually exhibit better average case performance than both the ellipsoid method and the method in [18], applicable when five or more receivers are available. The ellipsoid method is easily extended to handle more than the minimal nine transmitters, but not easily extended to handle more than the minimal four receivers. Although none of the methods are formally optimal in any sense, they are in closed form, fast, and can serve as initializations for further nonlinear optimization if need be. Both methods are significantly faster than the method in [18], and the matrix factorization method performs better in reconstructing the receiver positions.

In a real world experiment in an indoor environment, both methods perform well and the matrix factorization method reconstructs microphone positions with an average of 5 cm error from the previous 14 cm in [18].

Future work of interest is developing a method for calibration of UTOA networks not using the far field assumption, thus being able to solve problems when the far field assumption is far from true.

Acknowledgements The research leading to these results has received funding from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) within the Research Training Group 1103 (Embedded Microsystems), the strategic research projects ELLIIT and eSSENCE, and Swedish Foundation for Strategic Research projects ENGROSS and VINST (grants no. RIT08-0075 and RIT08-0043).

References

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

[2] A. Cirillo, R. Parisi, and A. Uncini, "Sound mapping in reverberant rooms by a robust direct method," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 31 2008-april 4 2008, pp. 285 –288.

[3] M. Cobos, A. Marti, and J.J. Lopez, "A modified srp-phat functional for robust real-time sound source localization with scalable spatial sampling," *Signal Processing Letters, IEEE*, vol. 18, no. 1, pp. 71 –74, jan. 2011.

[4] Hoang Do, H.F. Silverman, and Ying Yu, "A real-time srp-phat source location implementation using stochastic region contraction(src) on a large-aperture microphone array," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, april 2007, vol. 1, pp. I–121 –I–124.

[5] S. Thrun, "Affine structure from sound," in *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, Cambridge, MA, 2005, MIT Press.

[6] Y. Kuang, E. Ask, S. Burgess, and K. Åström, "Understanding toa and tdoa network calibration using far field approximation as initial estimate," in *ICPRAM*, 2012.

[7] R. Biswas and S. Thrun, "A distributed approach to passive localization for sensor networks," in *Proceedings of the National Conference on Artificial Intelligence*. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2005, vol. 20, p. 1248.

[8] P. Pertila, M. Mieskolainen, and M.S. Hamalainen, "Passive self-localization of microphones using ambient sounds," in *2012 Proceedings of the 20th*

European Signal Processing Conference (EUSIPCO). IEEE, 2012, pp. 1314–1318.

[9] C. Schindelhauer, Z. Lotker, and J. Wendeberg, "Network synchronization and localization based on stolen signals," in *Proceedings of 18th International Colloquium on Structural Information and Communication Complexity (SIROCCO)*, 2011.

[10] T. Janson, C. Schindelhauer, and J. Wendeberg, "Self-localization application for iphone using only ambient sound signals," in *Proceedings of the 2010 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Nov. 2010, pp. 259–268.

[11] Zheng Sun, Aveek Purohit, Kaifei Chen, Shijia Pan, Trevor Pering, and Pei Zhang, "Pandaa: Physical arrangement detection of networked devices through ambient-sound awareness," in *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp)*. ACM, 2011, pp. 425–434.

[12] Yubin Kuang, Simon Burgess, Anna Torstensson, and Kalle Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *Proc. of ICASSP*, 2013.

[13] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, APR 2005.

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

[15] Y. Kuang and K. Åström, "Stratified sensor network self-calibration from tdoa measurements," in *EUSIPCO*, 2013.

[16] R. Biswas and S. Thrun, "A passive approach to sensor network localization," in *IROS 2004*, 2004.

[17] Johannes Wendeberg, Fabian Höflinger, Christian Schindelhauer, and Leonhard Reindl, "Calibration-free tdoa self-localization," *Journal of Location Based Services*, 2013.

[18] S. Burgess, Y. Kuang, and K. Åström, "Node localization in unsynchronized time of arrival sensor networks," in *21st International Conference on Pattern Recognition (ICPR 2012), Proceedings of*. 2012, pp. 2042–2046, International Association for Pattern Recognition (IAPR) & IEEE.

[19] Martin A. Fischler and Robert C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, jun 1981.

[20] J. Wendeberg, T. Janson, and C. Schindelhauer, "Self-localization based on ambient signals," *Theoretical Computer Science*, vol. 453, pp. 98–109, 2011.

[21] Nikolai Nawri, "Berechnung von kovarianzellipsen," Received from: `http://imkbemu.physik.uni-karlsruhe.de/~eisatlas/covariance_ellipses.pdf`, 1996.

[22] S. T. Birchfield and A.. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *IEEE transactions on Speech and Audio Processing*, vol. 13, no. 5, 2005.

[23] T.S. Huang K.S. Arun and S.D. Blostein, "Least-squares fitting of two 3-d point sets," *IEEE Transactions on pattern analysis and machine intelligence*, 1987.

D

Paper D

TOA Based Self-Calibration of Dual Microphone Array

Zhayida Simayijiang, Simon Burgess, Yubin Kuang, and Kalle Åström

Centre for Mathematical Sciences, Lund University, Lund, Sweden

Abstract

In this paper we study the TOA (time-of-arrival) based self-calibration problem of dual microphone array for known and unknown rack distance, and also for affine space with different dimensions for receiver and sender spaces. Particularly we analyze the minimum cases and present minimum solvers for the case of microphones and speakers in 3D/3D, in 2D/3D, and in 3D/2D, with given or unknown rack length. We identify for each of these minimal problems the number of solutions in general and develop efficient and numerically stable, non-iterative solvers. Solving these problems are of both theoretical and practical interest. This includes understanding what the minimal problems are and how and when they can be solved. The solvers can be used to initialize local optimization algorithms for finding the maximum likelihood estimate of the parameters. The solvers can also be used for robust estimation of the parameters in the presence of outliers, using e.g. RANSAC algorithms. We demonstrate that the proposed solvers are numerically stable in synthetic experiments. We also demonstrate how the solvers can be used with the RANSAC paradigm. We also apply our method for several real data experiments, using ultra-wide-band measurements and using acoustic data.

Key words: Time-of-Arrival, Dual Microphone Array, Self-Calibration, Minimal Solver.

1 Introduction

The problem of sensor (or microphone) localization has broad interest in general. Time-of-arrival (TOA) and time-difference-of-arrival (TDOA) measurements are used in applications ranging from radio based positioning to beam-forming and audio sensing. In a more general setting, this is an inverse problem, where one tries to determine positions x_i using distance measurements $d_{i,j} = |x_i - x_j|$, [1]. As such these problems are applicable to a large range of applications such as (i) node determination using signal strength measurements of e.g. wifi or bluetooth, (ii) node localization using ultra-wide-band distance measurements, (iii) node localization using phase based radio measurements or (iv) node localization using acoustic measurements. In this more general setting there are also applications to node calibration of sensor networks, and also to 3D atom positioning determination using NMR [2]. Although such problems have been studied extensively in the literature in the form of localization of e.g. a sound source using a calibrated detector array, see e.g. [3–6], the problem of calibration of a sensor array using only measurement, the initialization problem for sensor network calibration, has received comparatively less attention, see [7, 8]

Antenna self-calibration techniques remain of great importance to many application beyond the microphone arrays. Several previous contributions addressing the self-calibration problem rely on prior knowledge or extra assumptions of locations of the sensors to initialize the problem. Some techniques are to manually measure the inter-distance between pairs of microphones and use multi-dimensional scaling [9] to compute receiver locations, [10], to use GPS [11] to get approximate locations, to use additional speakers close to the receivers [12–14], or to use a partially calibrated receiver array [15]. Being able to solve the self-calibration problem without using more than the TOA measurements between speakers and receivers open up for interesting applications. Iterative methods exist for TOA or TDOA based self-calibration [16, 17]. However, such methods are dependent on initialization and can get stuck in local minima. For a general graph structure, one can relax the TOA-based calibration problem as a semi-definite program [18]. Initialization of TOA sensor networks using only measurements has been studied in [19, 20], where solutions to the minimal cases of three senders and three receivers in the plane, or six senders and four receivers in 3D are given. Initialization of TDOA networks is studied in [21] and refined in [22] where a solution to non-minimal case of 9 receivers and 4 speakers in 3D was derived. In [23] a far field approximation was utilized to calibrate both

Figure 1: The problem of inferring structure from distance measurements, is relevant for measurements using audio and radio. The top part of the image contains one of the microphones (The t.bone MM1) and one of the speakers used for acoustic experiments. The bottom part of the image contains one ultra-wide-band tag and a ultra-wide-band equipped mobile phone used for radio experiments.

TOA and TDOA networks. In [17, 24], algorithms for far-field unsynchronized receivers were also proposed. In the far-field approximation, it is assumed that the distances between the speakers and receivers is considerably larger than between receivers. [19–24] attempt to solve the self-calibration problem with either minimal or close to minimal data. Studying minimal cases is both of theoretical importance and essential to develop fast stable algorithms suitable in random sample consensus (RANSAC) [25] schemes. RANSAC has the advantage of doing parameter estimation and at the same time weed out outliers in noisy or incorrectly matched data, which is a common problem in TOA/TDOA applications.

In this paper we focus on the previously unsolved problem of finding positions of a set of receivers and speakers, where pairs of receivers are set on a rigid rack, using only time-of-arrival (TOA) measurements between receivers and speakers with unknown positions, see Fig.2. We show in what constellations of receivers and speakers the self-calibration problem has a solution, and present numerically stable closed form solvers for these minimal cases. Applications can be in robotics and SLAM, where a robot is equipped with stereo microphones in a rigid constellation, moving through an environment with a number of fixed loudspeakers in unknown positions. To here be able to lock in part of the map, i.e. find

loudspeaker positions, and simultaneously figure out movement is a crucial step in SLAM, cf. [26]. This corresponds to the dual receiver rack self-calibration we study here. Minimal algorithms can further be used in a RANSAC setting to weed out outliers and identify parts of the measurement set which is good. An alternative could be that there is a set of fixed microphones in unknown position in the room and that the robot has a pair of loudspeakers attached to it. Recently mobile devices e.g. iPhone 5s also come equipped with dual microphones. Thus, an application could be to figure out movement of the mobile devices, using TDOA measurements from unknown sources. Solving the corresponding TDOA calibration problem often involves a two step process: First, figuring out the offsets and then solving the TOA calibration problem, see e.g. [22, 27]. Another application is that of ultra-wide-band distance measurements, e g such that measure distances to tags, see Figure 9. Here the tag calibration and the movement of a rack of two such receivers could be determined using the distance measurements only. The dual receiver rack self-calibration we study here also has the advantage of needing fewer measurements than the corresponding self-calibration problems for unconstrained receivers, and is thus better suited in RANSAC schemes where the setting applies. In addition, the constraint that the microphones are on a rack of the same length is explicitly enforced.

2 The TOA-based Microphone-rack Calibration Problem

We study the TOA-based mic-rack calibration problem for dual microphone arrays. A dual microphone array is a rigid array with two receivers and we set all racks to have same length c between receivers. The problem setting can be seen in Fig.2, where the microphone-racks can either be different, or one rack moving between measurement positions. We describe the problem of locating signal emitters and receivers, where pairs of receivers are fixed on a rack. It is considered further if the receicver distance on the rack is given ('calibrated' rack) or must be found. Measuring the absolute distance between every signal and receiver constrains the possible locations, which also known as Time-of-Arrival (TOA).

We assume that (i) the speed of the medium, v, is known, and thus all time measurements are transformed to distances by multiplication by v, (ii) receivers can distinguish which TOA signal comes from which transmitter. This can be done in practice by e.g. separating the signals temporally or by frequency. The

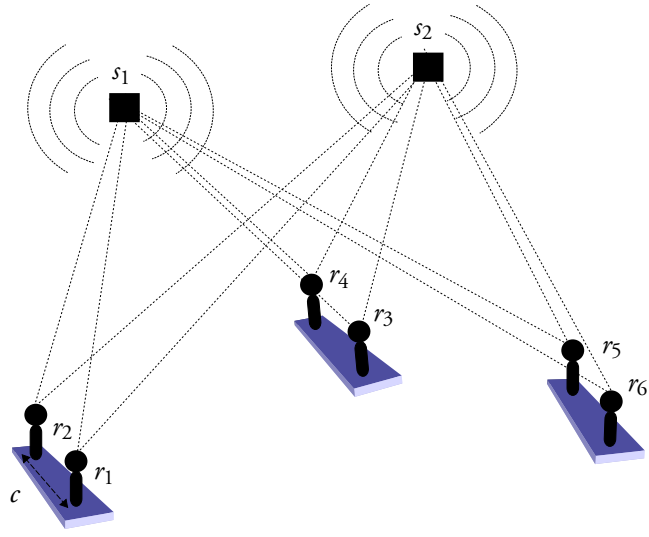


Figure 2: Sketch of setting. Rigid pair of receivers with unknown positions get TOA signals from unknown sender positions. The paper addresses the problem of recovering the displacement of a set of microphones and sound source positions based on the TOA measurements between microphones and sources. In particular, microphones are organized in couples, with a fixed distance c between each microphone of the couple. Here $r_1, \ldots, r_6$ are receivers and s_1, s_2 transmitters, respectively.

TOA dual rack calibration problem can then be defined as follows.

Let m be the number of microphone racks, k_r be the dimension of affine space of microphone racks, n be the number of sound sources and k_s be the dimension of affine space of sound sources. Let $\mathbf{r}_i$, $i = 1, \ldots, 2m$, be the spatial coordinates of $2m$ receivers (e.g. microphones) and $\mathbf{s}_j$, $j = 1, \ldots, n$, be the spatial coordinates of n speakers (e.g. sound events), and $c = ||\mathbf{r}_{2k} - \mathbf{r}_{2k-1}||_2$, for $k = 1, \ldots, m$, is fixed length of the rack. We have two problems of concern:

Problem 2.1. (Calibrated) Given absolute distance measurements $d_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2$ and a fixed known length between receiver pairs on the same rack c, determine receiver positions $\mathbf{r}_i$ and transmitter positions $\mathbf{s}_j$.

Problem 2.2. (Uncalibrated) Given absolute distance measurements

$d_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2$, determine receiver positions $\mathbf{r}_i$ and transmitter positions $\mathbf{s}_j$, as well the constant length of the receiver racks c.

One can only hope to recover a solution up to rotation/mirroring and translation of coordinate system, as any such transformation applied to a solution $\mathbf{r}_i, \mathbf{s}_j$ result in the same measurements d_{ij}.

In the following, we called the problem with known/unknown rack length c as *calibrated/uncalibrated* mic-rack self-calibration problem.

2.1 Identifying Minimal Problems

Depending on the number of receiver racks and speakers, we first characterize when such problems are well-defined i.e. when there is a finite positive number of solutions to the problem. We are particularly interested in the minimal problems where minimal numbers of receiver racks and speakers are required to solve the problem. One way to identify such minimal problems is to study the degrees of freedom of the problems F and the number of measurements M. The necessary condition for a problem to be minimal is that $F = M$ (comparing number of measurements to the degrees of freedom, we see if they are equal). For instance, for the case where the m receiver racks and the n speakers both span a 3D affine space, we have $F = 6m + 3n - 6$ (here -6 takes care of the *gauge freedom* i.e. rotation and translation ambiguity in the reconstruction). The condition that $F = M$ is the necessary condition, but not sufficient condition for identifiability of the problem, which can be seen in 2 racks on 2D and 4 speakers on 3D case, see Section 3.2.3. There are in some sense fewer measurements than $2mn + m$ due to the constraints imposed by the problem for these constellations.

If we denote $k_d = \min(k_r, k_s)$, then $q_d = k_d(k_d+1)/2$ is the *gauge freedom* of the d-dimensional space (we can only determine the positions up to an unknown rigid transforms i.e. rotation and translation). Specifically, we have for calibrated cases (known rack length c):

$$2k_r m + k_s n - q_d = 2mn + m. \tag{1}$$

and uncalibrated cases (unknown rack length c):

$$2k_r m + k_s n - q_d = 2mn + m - 1. \tag{2}$$

Based on (1) and (2), we have identified the following potential minimal problems: we have $M = 2mn + m$ and $M = 2mn + m - 1$ for calibrated and uncalibrated cases, respectively. By finding m, n such that $F = M$, we can identify

m\n	2	3	4	5
2	-	$G_{2,3}$	$(G_{3,3}, G^u_{2,3}, G^u_{3,2})$	$G^u_{3,3}$
3	$G_{2,3}$	$G_{3,2}$	-	-
4	$G^u_{2,3}$	$G^u_{3,2}$	-	-

Table 1: Here m and n represents number of racks and speakers respectively. G_{k_m,k_n} denotes the corresponding case in the table has degrees of freedom equal to number of unknowns for calibrated dual microphone racks in dimension k_m and sounds in dimension k_n. $G^u_{k_m,k_n}$ denotes the uncalibrated case of G_{k_m,k_n}. For ease of readability, different colors used for different dimensions, which are black for 3D-3D cases, red for 2D-3D cases, blue for 3D-2D cases. Note: case of 1 rack and 1 speaker (calibrated) is minimal case and it corresponds to TOA trilateration in 2D, which is not put into the table.

potential minimal cases. With this type of analysis, we have identified a set of potential minimal problems. These cases are summarized in Table 1, however, this by itself is not a sufficient condition for minimality. The cases in Table 1 are only based on calculations of degrees of freedom and number of constraints. Some of the 'minimal cases' might not be solvable. A further investigation of the different problems is needed.

2.2 Problem Formulation

The goal is to derive a low-degree polynomial equations system with few unknowns. We start by deriving a set of new equations. Since the distance measurements are assumed to be real and positive one does not lose any information by squaring them, i.e.

$$d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j) = \mathbf{r}_i^T\mathbf{r}_i + \mathbf{s}_j^T\mathbf{s}_j - 2\mathbf{r}_i^T\mathbf{s}_j.$$

The problem is significantly easier to analyze and solve by forming new equations according to the following linear combinations of d_{ij}^2. Given the TOA measurements between receivers and the speakers, we have $2mn$ equations in the following four types:

(A) 1 equation $d_{11}^2 = (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1)$.

(B) $n-1$ equations of type $d_{1j}^2 - d_{11}^2 = \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 - 2\mathbf{r}_1^T(\mathbf{s}_j - \mathbf{s}_1)$, for $j = 2, \ldots, n$.

(C) $2m-1$ equations of type $d_{i1}^2 - d_{11}^2 = \mathbf{r}_i^T\mathbf{r}_i - \mathbf{r}_1^T\mathbf{r}_1 - 2(\mathbf{r}_i - \mathbf{r}_1)^T\mathbf{s}_1$, for $i = 2, \ldots, 2m$.

(D) $(2m-1)(n-1)$ equations of type $d_{ij}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2 = -2(\mathbf{r}_i - \mathbf{r}_1)^T(\mathbf{s}_j - \mathbf{s}_1)$, for $i = 2, \ldots, 2m, j = 2, \ldots, n$.

And we also could have following equations for known or unknown rack distance, respectively:

(E) m equations of type $c_{2k-1,2k}^2 = (\mathbf{r}_{2k-1} - \mathbf{r}_{2k})^T(\mathbf{r}_{2k-1} - \mathbf{r}_{2k})$, for $k = 1, \ldots, m$.

(E′) $m-1$ equations of type $||\mathbf{r}_1 - \mathbf{r}_2|| - ||\mathbf{r}_{2k-1} - \mathbf{r}_{2k}|| = 0$, for $k = 2, \ldots, m$

Using these equations, we now describe methods to solve the polynomial systems for different minimal problems. The main steps of the method includes:

- Choosing parametrizations and rewriting equations.
- Factorization.
- Solving for the unknown transformation.

3 Solving Minimal Problems

The solution strategy is to use a factorization step first to reconstruct the positions up to an unknown affine transformation $\mathbf{L}$ and $\mathbf{b}$. By collecting terms, the equations of type (D) can be written in matrix form $\tilde{\mathbf{D}} = \mathbf{r}^T\mathbf{s}$ with $\mathbf{r}_i - \mathbf{r}_1$ as columns of $\mathbf{r}$ and $-2\left(\mathbf{s}_j - \mathbf{s}_1\right)$ as columns of $\mathbf{s}$. The rank of $\tilde{\mathbf{D}}$ depends on the dimensionality of the affine span of the receivers and the speakers respectively. For instance, if we assume that both of the racks and the speakers are in 3D, then the matrix $\tilde{\mathbf{D}}$ also has rank 3. By factorizing $\tilde{\mathbf{D}}$ which is of rank 3 using e.g. singular value decomposition, we can compute the vectors to all receivers and speakers from unknown initial/reference positions ($\mathbf{r}_1$ and $\mathbf{s}_1$) up to an unknown full-rank 3×3 transformation $\mathbf{L}$ such that $\tilde{\mathbf{D}} = \tilde{\mathbf{r}}^T\mathbf{L}^{-1}\mathbf{L}\tilde{\mathbf{s}} = \mathbf{r}^T\mathbf{s}$. Depending on how one fixes the gauge freedom, the unknown $\mathbf{b}$ enters the equations in different ways.

By a good choice of parametrization of the problems it can be shown that the equations of Types C, E (or E′) are linear in the unknowns and the equations of Types A and B can be used to form polynomial equations.

3.1 3D-racks and 3D-Speakers

In this section, we solve the minimal problems for the cases where both racks and speakers are in 3D.

To solve for the unknown transformation and reference positions, we now utilize the nonlinear constraints in equations of Type A, B and C. First we can fix the gauge freedom by choosing the location $\mathbf{r}_1$ at the origin and also parameterize $\mathbf{s}_1$ as $\mathbf{Lb}$ where $\mathbf{b}$ is a 3×1 vector. Given that $\mathbf{r} = \mathbf{L}^{-T}\tilde{\mathbf{r}}$ and $\mathbf{s} = \mathbf{L}\tilde{\mathbf{s}}$. This gives

$$\begin{aligned} \mathbf{r}_i &= \mathbf{L}^{-T}\tilde{\mathbf{r}}_i, \ i = 2 \dots 2m \\ \mathbf{s}_j &= \mathbf{L}(\tilde{\mathbf{s}}_j^* + \mathbf{b}), \ j = 2 \dots n, \end{aligned} \tag{3}$$

where $\tilde{\mathbf{s}}_j^* = \tilde{\mathbf{s}}_j/(-2)$. Using the parametrization above and also letting $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$ the equations of type (A), (B), (C) and (E) become

$$d_{11}^2 = \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}, \tag{4}$$

$$d_{1j}^2 - d_{11}^2 = \tilde{\mathbf{s}}_j^{*T}\mathbf{H}^{-1}\tilde{\mathbf{s}}_j^* + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{s}}_j^*, \tag{5}$$

$$d_{i1}^2 - d_{11}^2 = \tilde{\mathbf{r}}_i^T\mathbf{H}\tilde{\mathbf{r}}_i - 2\mathbf{b}^T\tilde{\mathbf{r}}_i, \tag{6}$$

$$c_{2k-1,2k}^2 = \begin{cases} \tilde{\mathbf{r}}_2^T\mathbf{H}\tilde{\mathbf{r}}_2, & k = 1 \\ \tilde{\mathbf{r}}_{2k-1}^T\mathbf{H}\tilde{\mathbf{r}}_{2k-1} - 2\tilde{\mathbf{r}}_{2k-1}^T\mathbf{H}\tilde{\mathbf{r}}_{2k} + \\ +\tilde{\mathbf{r}}_{2k}^T\mathbf{H}\tilde{\mathbf{r}}_{2k}, & k > 1. \end{cases} \tag{7}$$

3.1.1 Case of 2 racks and 4 speakers (calibrated)

There are in total 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively). By utilizing $\mathbf{H}^{-1} = \text{adj}(\mathbf{H})/\det(\mathbf{H})$, where $\text{adj}(\mathbf{H})$ is the adjoint of $\mathbf{H}$, we can multiply equations in (4)-(5) by $\det(\mathbf{H})$ to rewrite them as polynomials equations. There are three linear equations of type C and two linear equation of type E. Since we have 5 linear equations, by linear elimination we can parameterize $\mathbf{H}$ and $\mathbf{b}$ in terms of $9 - 5 = 4$ unknowns $\mathbf{x} = (x_1, x_2, x_3, x_4)$, therefor they become

$\mathbf{H} = \mathbf{H}(x_1, x_2, x_3, x_4)$, $\mathbf{b} = \mathbf{b}(x_1, x_2, x_3, x_4)$. After multiplying $\det(\mathbf{H})$ on both sides of Type (A) and (B) equations, we now obtain polynomial system with four equations:

$$\det(\mathbf{H})d_{11}^2 = \mathbf{b}^T \operatorname{adj}(\mathbf{H})\mathbf{b} \tag{8}$$
$$\det(\mathbf{H})(d_{12}^2 - d_{11}^2) = \tilde{\mathbf{s}}_2^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{s}}_2^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{s}}_2^* \tag{9}$$
$$\det(\mathbf{H})(d_{13}^2 - d_{11}^2) = \tilde{\mathbf{s}}_3^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{s}}_3^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{s}}_3^* \tag{10}$$
$$\det(\mathbf{H})(d_{14}^2 - d_{11}^2) = \tilde{\mathbf{s}}_4^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{s}}_4^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{s}}_4^* \tag{11}$$

in the four unknowns. While multiplying $\det(\mathbf{H})$ introduces false solutions, we utilize the same saturation technique as in [20] to remove such solutions. Using algebraic geometric tools, we verify that this system has in general 12 solutions. We then solve this system with efficient polynomials solvers based on [28]. After solving for $\mathbf{H}$, L can be calculated with Cholesky factorization. $\mathbf{L}$ is thus only determined up to a matrix R where $R^T R = I$, which coincides with the gauge freedom of rotating and/or mirroring our solution. Some of the solutions obtained are complex. Some solutions, although real, give invalid matrices $\mathbf{H}$ that are not positive definite which does not have a real decomposition into $(\mathbf{L}^T\mathbf{L})^{-1}$ and does not produce real solutions.

3.1.2 Case of 2 racks and 5 speakers (uncalibrated)

For this case, there are 9 equations (1 of Type A, 4 of Type B, 3 of Type C and 1 of Type E′) and 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively). We follow a similar solution scheme as for the case of 2 racks and 4 speakers case. By linear elimination using the 4 linear constraints of type C and E, we can express $\mathbf{H}$ and $\mathbf{b}$ in terms of $9 - 4 = 5$ unknowns $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$. The remaining five constraints (1 of Type A, and 4 of Type B) gives a polynomial system with 28 solutions after a saturation procedure similar to the previous case. Again we use the scheme in [28] to produce a numerically stable and efficient solution.

3.2 2D-racks and 3D-speakers

In this section, we solve the minimal problems for cases where the racks in 2D and speakers are in 3D.

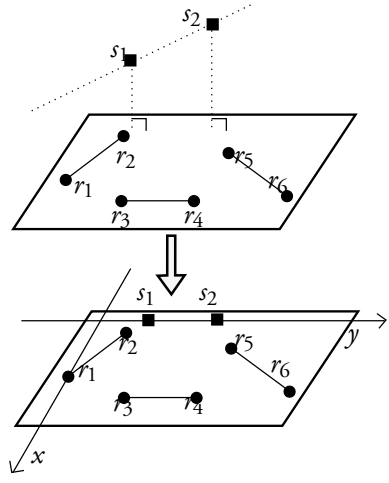


Figure 3: Three microphone racks are on 2D plane and two speakers are on 3D space, and get projection of speaker positions on 2D plane.

3.2.1 Case of 3 racks and 2 speakers (calibrated)

As shown in Figure. 3, we get orthogonal projection of $\mathbf{s}_1$ and $\mathbf{s}_2$ on the plane which microphone racks lie. We set

$$\mathbf{r} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ 0 & r_{22} & r_{23} & r_{24} & r_{25} & r_{26} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mathbf{s} = \begin{bmatrix} 0 & 0 \\ s_{21} & s_{22} \\ s_{31} & s_{32} \end{bmatrix}$$

By factorizing $\tilde{\mathbf{D}}$ which is of rank 1 in this case, we can compute all receivers and speakers from unknown initial/reference positions up to the unknown transformation $L = l$, which is now only an unknown constant. We now utilize the nonlinear constraints in equations of Type A, B, C and E.

We know that one can reconstruct for $\mathbf{r}$ and $\mathbf{s}$ with factorization up to unknown transformation l such that $\tilde{\mathbf{D}} = \tilde{\mathbf{r}}^T \frac{1}{l} l \tilde{\mathbf{s}} = \mathbf{r}^T \mathbf{s}$, we put $\tilde{\mathbf{r}} = [0 \quad \tilde{\mathbf{D}}^T]$ and

$\tilde{\mathbf{s}} = [0 \quad \mathbf{1}]$. We fix the gauge freedom by choosing the location $\mathbf{r}_1 = [r_{11}, 0, 0]^T$, $\mathbf{r}_i = [r_{1i}, r_{2i}, 0]^T, i = 2, \ldots, 6$ and $\mathbf{s}_j = [0, s_{2j}, s_{3j}]^T, j = 1, 2$, and also denote second row of $\mathbf{r}$ and $\mathbf{s}$ as $\overline{\mathbf{r}}$ and $\overline{\mathbf{s}}$, respectively. This gives

$$\begin{aligned} \overline{\mathbf{r}}_i &= \frac{1}{l}\tilde{\mathbf{r}}_i, \ i = 2, \ldots, 6 \\ \overline{\mathbf{s}}_j &= l(\tilde{\mathbf{s}}_j + b)/(-2), \ j = 1, 2. \end{aligned} \tag{12}$$

The equations of type (A), (B), (C) and (E) then become

$$d_{11}^2 = r_{11}^2 + \frac{l^2 b^2}{4} + s_{31}^2, \tag{13}$$

$$d_{12}^2 - d_{11}^2 = \frac{l^2}{4}\tilde{\mathbf{s}}_2^2 + \frac{l^2 b}{2}\tilde{\mathbf{s}}_2 + s_{32}^2 - s_{31}^2, \tag{14}$$

$$d_{i1}^2 - d_{11}^2 = r_{1i}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_i + b\tilde{\mathbf{r}}_i - r_{11}^2, i = 2, \ldots, 6 \tag{15}$$

$$c_{2k-1,2k}^2 = \begin{cases} r_{11}^2 + r_{12}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_2^2 - 2r_{11}r_{12}, & k = 1 \\ r_{1,2k-1}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_{2k-1} + r_{1,2k}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_{2k} - \\ \quad -2r_{1,2k-1}r_{1,2k} - \frac{2}{l^2}\tilde{\mathbf{r}}_{2k-1}\tilde{\mathbf{r}}_{2k}, & k = 2, 3. \end{cases} \tag{16}$$

There are in total 10 unknowns (1 for l, 1 for b, 6 for r_{1i} and 2 for s_{3j}). We have 10 equations (1 of Type A, 1 of Type B, 5 of Type C and 3 of Type E). By using the parametrization $\mathbf{x} = [h, b, v_1, v_2, v_3, v_4, v_5, v_6, u_{12}, u_{34}, u_{56}]^T$, where $h = \frac{1}{l^2}$, $v_i = r_{1i}^2, i = 1, \ldots, 6$ and $u_{2k-1,2k} = r_{1,2k-1}r_{1,2k}, k = 1, 2, 3$. We have 8 linear equations from (15) and (16), thus we can express all the unknowns in $\mathbf{x}$ linearly in terms of $[h, b, v_1]$:

$$\mathbf{A} = \begin{bmatrix} \tilde{\mathbf{r}}_2^2 & \tilde{\mathbf{r}}_2 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \tilde{\mathbf{r}}_3^2 & \tilde{\mathbf{r}}_3 & -1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \tilde{\mathbf{r}}_4^2 & \tilde{\mathbf{r}}_4 & -1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \tilde{\mathbf{r}}_5^2 & \tilde{\mathbf{r}}_5 & -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \tilde{\mathbf{r}}_6^2 & \tilde{\mathbf{r}}_6 & -1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \tilde{\mathbf{r}}_2^2 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & -2 & 0 & 0 \\ \tilde{\mathbf{r}}_4^2 - 2\tilde{\mathbf{r}}_3\tilde{\mathbf{r}}_4 + \tilde{\mathbf{r}}_3^2 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & -2 & 0 \\ \tilde{\mathbf{r}}_5^2 - 2\tilde{\mathbf{r}}_5\tilde{\mathbf{r}}_6 + \tilde{\mathbf{r}}_6^2 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & -2 \end{bmatrix}$$

Algorithm 1 Solver for 3 racks and 2 speaker, 2D-3D, calibrated case

Input: Distances matrix $\mathbf{D}$.
Output: Microphone and sound position.

1: Square all distances in $\mathbf{D}$.
2: Form the four types of distance squared differences.
3: Use d_{ij} to determine $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{s}}$.
4: Get linear equations in $x = [h, b, v_1, v_2, v_3, v_4, v_5, v_6, u_{12}, u_{34}, u_{56}]^T$.
5: Set equations type C and also equations from rack constraints.
6: Express $[v_2, v_3, v_4, v_5, v_6, u_{12}, u_{34}, u_{56}]$ linearly in terms of $[h, b, v_1]$.
7: Form quadratic equations $u_{12}^2 = v_1 v_2, u_{34}^2 = v_3 v_4, u_{56}^2 = v_5 v_6$.
8: Solve polynomial equations using action matrix techniques. (We get two possible solutions, maybe only one is feasible.)
9: **for** each solution **do**
10: Backtrack and calculate all unknown parameters.
11: Test the signs of the pairs of $\mathbf{r}$ for each rack.
12: **end for**

Let $\mathbf{A} = [\mathbf{A}_1 \quad \mathbf{A}_2]$ where $\mathbf{A}_1$ is first 3 columns and $\mathbf{A}_2$ is last 8 columns of $\mathbf{A}$ and denote first 3 elements of $\mathbf{x}$ as $\mathbf{x}_1 = [h, b, v_1]$ and last 8 elements as $\mathbf{x}_2$, then we can construct the following equations:

$$[\mathbf{A}_1 \quad \mathbf{A}_2][\mathbf{x}_1 \quad \mathbf{x}_2]^T = \mathbf{A}_1\mathbf{x}_1 + \mathbf{A}_2\mathbf{x}_2 = \mathbf{B} \tag{17}$$

We could have a solution $\mathbf{x}_2 = \mathbf{A}_2^{-1}\mathbf{B} - \mathbf{A}_2^{-1}\mathbf{A}_1\mathbf{x}_1$. We then proceed to solve the three equations

$$u_{2k-1,2k}^2 = r_{1,2k-1}^2 r_{1,2k}^2, \quad k = 1, 2, 3. \tag{18}$$

in the three unknowns $[h, b, v_1]$ using the techniques in [28]. Resubstitution gives us the coordinated of $\mathbf{r}$ and $\mathbf{s}$. In general there are 2 solutions.

3.2.2 Case of 4 racks and 2 speakers (uncalibrated)

Similar to the previous calibrated case, we now have 12 equations (1 of Type A, 1 of Type B, 7 of Type C and 3 of Type E′). We again need to select ten columns of $\mathbf{A}$ such that $\mathbf{A}_2$ has full rank and low condition number. We solve the

problem in the same manner as the calibrated 3 racks and 2 speakers case. For the corresponding polynomial system, there is in general 6 solutions.

3.2.3 Case of 2 racks and 4 speakers (uncalibrated)

This case is actually under-determined even though it satisfies $F = M$. One way to explain this is the following. Adding one transmitter seems to give 4 measurements and 3 unknowns (unknown speaker positions in 3D), thus one obtains an additional constraint which could indicate that we can use that to solve the rack length c. But in fact we only get 3 (linearly independent) measurements due to the rank constraints on $\tilde{\mathbf{D}}$. Thus it is unsolvable.

3.2.4 Case of 2 racks and 3 speakers (calibrated)

We can here parameterize the two racks to be on the z-plane. Then we know that the matrix $\tilde{\mathbf{D}}$ is of rank-2, and $\mathbf{H}$ is a symmetric 2 by 2 matrix, and $\mathbf{b}$ is a 2 by 1 matrix. There are three linear equations of type C and two linear equations of type E and 5 unknowns (3 for $\mathbf{H}$ and 2 for $\mathbf{b}$). Thus we can solve for $\mathbf{H}$ and $\mathbf{b}$ linearly and resolve for the positions of $\mathbf{r}$ and the projection of $\mathbf{s}$ onto the z plane. Then we solve the z coordinates for the speakers simply using the distance measurements.

3.3 3D-racks and 2D-speakers

In this section, we solve the minimal problems for cases where the racks are in 3D and speakers are in 2D.

3.3.1 Case of 3 racks and 3 speakers (calibrated)

As Figure 4 shows, we carefully choose the gauge freedom so that microphone racks and speaker positions are of the following form:

$$\mathbf{r} = \begin{bmatrix} 0 & r_{12} & r_{13} & r_{14} & r_{15} & r_{16} \\ 0 & 0 & r_{23} & r_{24} & r_{25} & r_{26} \\ r_{31} & r_{32} & r_{33} & r_{34} & r_{35} & r_{36} \end{bmatrix}$$

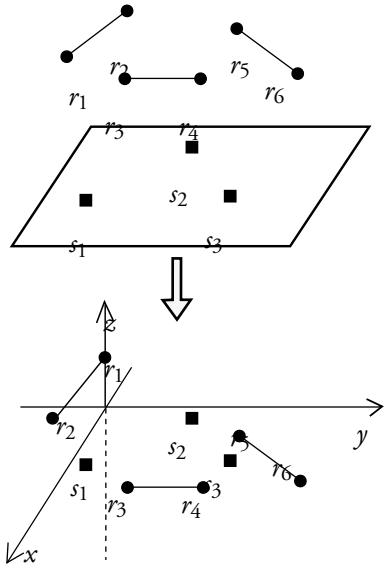


Figure 4: Three microphone racks are in 3D and three speakers in a 2D plane. The projections of microphone positions on the 2D plane are used.

$$\mathbf{s} = \begin{bmatrix} s_{11} & s_{12} & s_{13} \\ s_{21} & s_{22} & s_{23} \\ 0 & 0 & 0 \end{bmatrix}.$$

The solution strategy follows a similar technique as 2D-3D case, but this time $\tilde{\mathbf{D}}$ has rank 2, and we have different characterization of $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{s}}$ as $\tilde{\mathbf{r}} = [\mathbf{0}_{2\times 1} \;\; \tilde{\mathbf{D}}^T]$ and $\tilde{\mathbf{s}} = [\mathbf{0}_{2\times 1} \;\; \mathbf{1}]$. If we fix the translational part of the gauge freedom by choosing the location $\mathbf{r}_1 = [0, 0, r_{31}]^T$, $\mathbf{r}_2 = [r_{12}, 0, r_{32}]^T$ and also denote first two rows of $\mathbf{r}$ and $\mathbf{s}$ as $\bar{\mathbf{r}}$ and $\bar{\mathbf{s}}$, respectively, then we get our equation system in a similar way.

Then $\bar{\mathbf{r}}_i$ and $\bar{\mathbf{s}}_j$ gives ith and jth column of $\bar{\mathbf{r}}$ and $\bar{\mathbf{s}}$, respectively.

$$\begin{aligned} &\bar{\mathbf{r}}_i = \mathbf{L}^{-T}\tilde{\mathbf{r}}_i, \; i = 2, 3 \ldots 6, \\ &\bar{\mathbf{s}}_j = \mathbf{L}(\tilde{\mathbf{s}}_j + \mathbf{b})/(-2), \; j = 1, 2, 3. \end{aligned} \tag{19}$$

Using this parametrization and also let $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$, the equations of type (A), (B), (C) and (E) can then be simplified as

$$d_{11}^2 = r_{31}^2 + \frac{\mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}}{4}, \tag{20}$$

$$d_{1j}^2 - d_{11}^2 = \frac{\tilde{\mathbf{s}}_j^T\mathbf{H}^{-1}\tilde{\mathbf{s}}_j + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{s}}_j}{4} \quad j = 2,3 \tag{21}$$

$$d_{i1}^2 - d_{11}^2 = \tilde{\mathbf{r}}_i^T\mathbf{H}\tilde{\mathbf{r}}_i + \mathbf{b}^T\tilde{\mathbf{r}}_i + r_{3i}^2 - r_{31}^2, \quad i = 2,\ldots,6 \tag{22}$$

$$c_{2k-1,2k}^2 = \begin{cases} r_{31}^2 + \tilde{\mathbf{r}}_{12}^T\mathbf{H}\tilde{\mathbf{r}}_{12} + r_{32}^2 - 2r_{31}r_{32}, & k = 1 \\ \\ (\tilde{\mathbf{r}}_{1,2k}^T\mathbf{H}\tilde{\mathbf{r}}_{1,2k} + \tilde{\mathbf{r}}_{2,2k}^T\mathbf{H}\tilde{\mathbf{r}}_{2,2k}) + \\ +(\tilde{\mathbf{r}}_{1,2k-1}^T\mathbf{H}\tilde{\mathbf{r}}_{1,2k-1} + \tilde{\mathbf{r}}_{2,2k-1}^T\mathbf{H}\tilde{\mathbf{r}}_{2,2k-1}) - \\ -2(\tilde{\mathbf{r}}_{1,2k}^T\mathbf{H}\tilde{\mathbf{r}}_{1,2k-1} + \tilde{\mathbf{r}}_{2,2k}^T\mathbf{H}\tilde{\mathbf{r}}_{2,2k-1}) + \\ +r_{3,2k}^2 + r_{3,2k-1}^2 - 2r_{3,2k-1}r_{3,2k}. & k = 2,3 \end{cases} \tag{23}$$

For this case, there are 11 equations (1 of Type A, 2 of Type B, 5 of Type C and 3 of Type E) with 8 of them linear, and 11 unknowns (6 for z-coordinate of receivers and 5 unknowns for $\mathbf{H}$ and $\mathbf{b}$). We use the same technique which is used in Section. 3.2.1. Using algebraic geometry tools, we verify that there are in general 16 solutions. By using the parametrization

$$\mathbf{x} = [H_{11}, H_{12}, H_{22}, b_1, b_2, v_1, v_2, v_3, v_4, v_5, v_6, u_{12}, u_{34}, u_{56}] \tag{24}$$

$$v_i = r_{3i}^2, i = 1,\ldots,6 \tag{25}$$

$$u_{2k-1,2k} = r_{3,2k-1}r_{3,2k}, k = 1,2,3. \tag{26}$$

We have 8 linear equations from (22) and (23), thus we can express all the unknowns in terms of $[H_{11}, H_{12}, H_{22}, b_1, b_2, v_1]$, similar to Section 3.2.1. We then proceed to solve equations (20), (21), and the three equations

$$u_{2k-1,2k}^2 = r_{3,2k-1}^2 r_{3,2k}^2, \quad k = 1,2,3. \tag{27}$$

in the unknowns $[H_{11}, H_{12}, H_{22}, b_1, b_2]$ using the techniques in [28].

3.3.2 Case of 4 racks and 3 speakers (uncalibrated)

We have 13 equations (1 of Type A, 2 of Type B, 7 of Type C and 3 of Type E′) and 13 unknowns (8 for z-coordinate of receivers and 5 unknowns for $\mathbf{H}$ and $\mathbf{b}$).

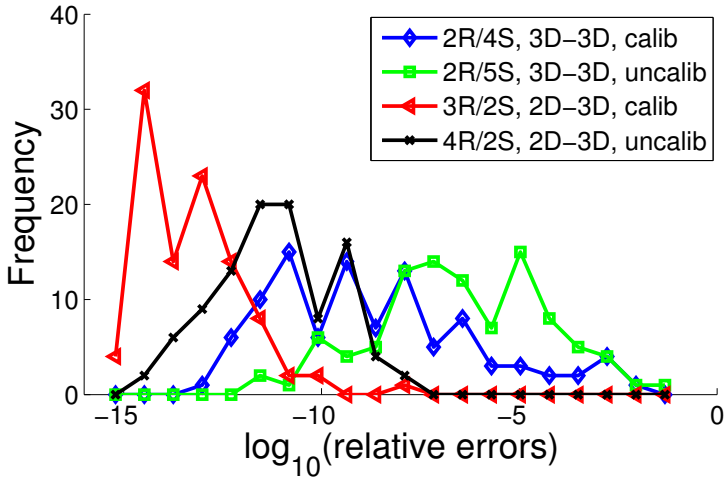


Figure 5: Histogram of relative errors for the solvers without additive Gaussian noise in 100 runs of solvers.

We follow the same strategy to parametrize the problem as the calibrated 3 racks and 3 speakers case. This problem is much more difficult and has in general 29 solutions.

3.3.3 Case of 2 racks and 4 speakers (uncalibrated)

We have 8 equations (1 of Type A, 3 of Type B, 3 of Type C and 1 of Type E′) and 9 unknowns (4 for z-coordinates of receivers and 5 unknowns for **H** and **b**). We do not have enough information to solve the case.

4 Using Minimal Solvers for Overdetermined Problems

The study of minimal configurations that are solvable are interesting from a theoretical viewpoint. By solving and understanding these minimal problems one obtains a better understanding of the geometrical problems and can identify possible failure cases for the estimation problem. The minimal solvers are also interesting from a practical viewpoints. Two main applications are (i) to obtain initial estimates for iterative methods for optimizing likelihood function and (ii) to remove outliers using sample and test algorithms such as RANSAC, cf. [25].

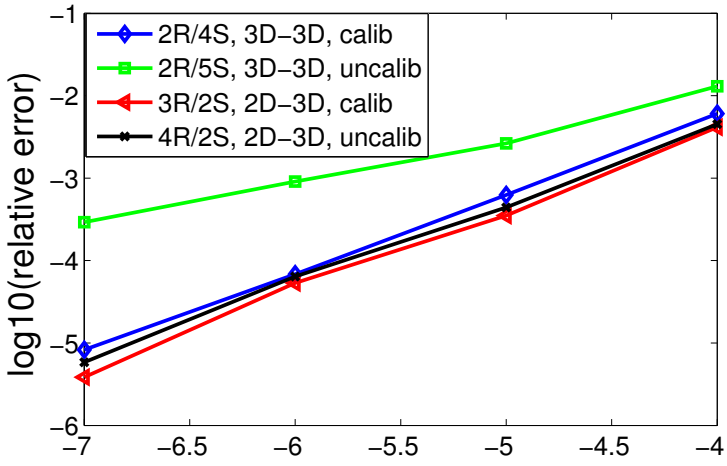


Figure 6: Mean of log10 relative errors for 100 runs, plotted against the standard deviation of the additive Gaussian noise.

4.1 Local optimization methods for estimating parameters for the overdetermined case

Assume as before that m is the number of microphone racks, k_r be the dimension of affine space of microphone racks, n be the number of sound sources and k_s be the dimension of affine space of sound sources. The unknown parameters are the spatial coordinates of the $2m$ receivers (e.g. microphones) $\mathbf{r}_i$, $i = 1, \ldots, 2m$, and the spatial coordinates of n speakers (e.g. sound events), $\mathbf{s}_j$, $j = 1, \ldots, n$ and possibly the unknown but constant distance between the receivers on the rack, c. In order to keep the presentation clear we will here only study the case where the rack length c is known. We are thus assuming that $||\mathbf{r}_{2k} - \mathbf{r}_{2k-1}|| = c$, for $k = 1, \ldots, m$. For brevity we will use z to denote the unknown parameter, i.e. $z = (\mathbf{r}_1, \ldots \mathbf{r}_m, \mathbf{s}_1, \ldots, \mathbf{s}_n)$. The measurements are the distance measurements $d_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2 + e_{ij}$. Here e_{ij} are used to denote the measurement errors. Assuming that these are independent, the maximum likelihood estimate of the parameter is found by maximizing

$$\max_z f(z) = \max_z \prod_{i,j} p(d_{ij}|\mathbf{r}_i, \mathbf{s}_j).$$

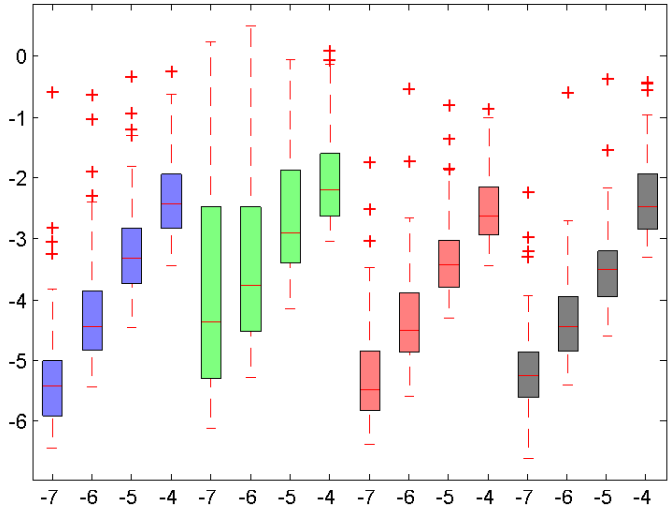


Figure 7: Box plot of mean reconstruction error. Legend is same with Figure. 6. It shows the shape of the distribution, its central value, and variability. This plot displays the full range of variation (from min to max), the likely range of variation, and the median, and also shows outliers.

One common choice for modeling the probability density functions p is to use the Gaussian distribution. By taking the negative logarithm this optimization problem becomes

$$\min_z f_{l_2}(z) = \sum_{i,j} (d_{ij} - |\mathbf{r}_i - \mathbf{s}_j)|^2).$$

The function f_{l_2} typically has many local minima. The fact that there are several solutions to some of the minimal problems as demonstrated above, is in fact a proof that there are such problems with several local minima. It is generally difficult to obtain good initial estimates of the parameters. However, once such initial estimates are obtained, it is relatively straightforward to improve the estimate by performing local optimization of f_{l_2} using e.g. Gauss-Newton method or any other local optimization method.

A disadvantage of optimizing f_{l_2} is that the solution is not only sensitive to

the initial estimate, but also sensitive to outliers in the data. Assuming Gaussian distribution is an implicit assumption that the probability of outliers is extremely small. For problems where there are more outliers it is often a better approach to assume a different distribution. One popular method is to minimize the one-norm, i.e. to minimize

$$\min_z f_{l_1}(z) = \sum_{i,j} |d_{ij} - |\mathbf{r}_i - \mathbf{s}_j)||.$$

This gives methods that are less sensitive to outliers. It is however still sensitive to the initial estimate of the parameter. Again this is because the function f_{l_1} also has many local minima.

In all of these cases it is possible to use the solutions from the minimal solvers to obtain initial estimates for non-linear optimization techniques.

4.2 Random sampling and case

Hypothesize and test ideas such as Random Sampling Consensus (RANSAC) can be quite effective for finding good parameter values in the presence of outliers. The main idea of the algorithm is to make a number of iterations, where in each iteration a minimal subset of the data is chosen. Even in the presence of outliers there might be a fair chance that the subset is outlier-free. This probability is higher the smaller the subset is. An algorithm for solving for the parameters using minimal data is then used. This estimated parameters are then tested on the remaining data. These tests can typically be done sequentially, so that each step is relatively robust to outliers and such that each test is computationally efficient. The idea here is that a good set of parameters will fit to a large portion of the remaining data.

5 Experiments

To be able to evaluate the quality of a solution, receivers and transmitter positions $\mathbf{r}_i$ and $\mathbf{s}_j$ are compared to ground truth positions $\mathbf{r}_{i,gt}$ and $\mathbf{s}_{j,gt}$. Positions are rotated, mirrored and translated so that the points are aligned using [29]. For comparing with computer vision reconstruction, the alignment of points is also done over scale. Relative errors are defined as $\|[\mathbf{r} - \mathbf{r}_{gt}\ \mathbf{s} - \mathbf{s}_{gt}]\|_F / \|[\mathbf{r}_{gt}\ \mathbf{s}_{gt}]\|_F$, where $\|\cdot\|_F$ is the Frobenius norm. For the additive noise $\overline{d_{ij}} = d_{ij} + \varepsilon, \varepsilon \in \mathcal{N}(0, \sigma)$,

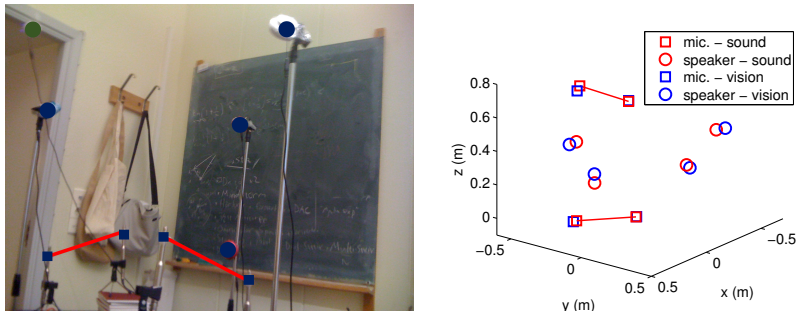
we use the Gaussian distribution.

Simulations were run for 100 times where ground truth receivers and transmitters where drawn uniformly over a unit cube around the origin. Half of the receivers were then fitted to a rack distance of 0.2 from their respective pairs, and measurements d_{ij} were created from ground truth. Relative errors for the solvers with and without additive Gaussian noise on the measurements can be seen in Figure 5 and Figure 6, respectively. Looking at Figure 5, the results of the minimal solver show that the algorithms has good numerical performance. For indicating whether a distribution is skewed and whether there are any unusual observations (outliers) in the mean error, we add standard deviation bar and outliers to the mean reconstruction error, which is showed in Fig. 7. The central rectangle spans the first quartile to the third quartile, a segment inside the rectangle shows the median and "whiskers" above and below the box show the locations of the minimum and maximum mean errors.

For testing the case of senders and receivers in 3D with a calibrated rack distance, an indoor experiment was carried out. A set of real data was obtained using four T-bone MM-1 microphones and four Roxcore portable speakers, connected to a Fast Track Ultra 8R sound card in an indoor environment, with speakers and microphones placed in an approximate $1.5 \times 1.5 \times 1.5\ m^3$ volume (Fig.8a).

TOA measurements were obtained by matching sounds from different speakers to sound flanks recorded from different microphones. The footprint of the matching was selected to be small as to avoid degeneration from reverberation effects. Synchronization was achieved by sharing the same local clock by having everything connected to the same sound card. The sounds were separated temporally so that the matching of which sound came from which speaker could be done. Matching was done using the beginning of emitted sounds, thus ignoring reflections as there exist a direct path between speakers and microphones. A reconstruction of the scene was made using computer vision techniques to be used as ground truth.

The reconstruction (Fig.8b) when compared to the vision-based reconstruction has an RMSE of 4.2 *cm* and 5.6 *cm* for microphones and speakers respectively. From Table 2, we can see that many minimal configuration have multiple solutions, the corrected solutions are always within the solutions we obtain, and there are false solutions. For these false solutions, we first remove the complex ones and further remove more by verifying using measurements from an additional microphone, seen in Fig. 8a as the microphone with a green dot on.



(a) (b)

Figure 8: Real microphone and speaker calibration setup. (a) the setup of microphones and speakers in an office. Blue square is for microphone position, and pair of mircophones are fixed on the same length rack which is shown as red line, and blue circle is for sound position. (b) the reconstructed sensor positions using TOA measurements (red) aligned with the positions estimated based on computer vision (blue).

Dimension	Cases	Nb. of solutions
3D-3D	2R/4S Calibrated	38
	2R/5S Uncalibrated	28
2D-3D	3R/2S Calibrated	2
	4R/2S Uncalibrated	6
	2R/4S Uncalibrated	unsolvable
	2R/3S Calibrated	
3D-2D	3R/3S Calibrated	16
	4R/3S Uncalibrated	29
	2R/4S Uncalibrated	underdetermined

Table 2: Number of solutions for all minimal cases of dual microphone array problem. Here R and S stands for rack and sender, respectively.

5.1 Overdetermined case with noise and outliers

To test how to use the new algorithms for overdetermined cases with outliers we have generated a bank (A) of test cases. For this test we have used 3 2-rigs moving in a plane and 40 sound sources. The sound sources are randomly genrated in a cube of size $2 \times 2 \times 2$, centered at the origin. The rigs are randomly generated in the plane $z = 0$, such that the rig distance has length 1. In total we generated 100 such random cases. For each case we calculate the distiances from each sound source to each microphone. We have thus 6×40 distance measurements for each case. In each case we randomly chose x% of the measurements to be outliers. To each of the measurements we added a random measurement error (Gaussian distribution with mean zero and standard deviation 0.001) if the measurement is considered to be an inlier and a larger error (rectangular distribution between 0.4 and 1.2) if the measurement is consedered to be an outlier.

To test the algorithms for the case of no outliers, a similar bank (B) of test sets were constructed without any outliers.

We test three algorithms for estimating the parameters,
$z = (\mathbf{r}_1, \dots \mathbf{r}_m, \mathbf{s}_1, \dots, \mathbf{s}_n)$.

1. Random initial estimate of z, followed by non-linear least squares optimization of the residuals. We repeat this process k times and choose the solution z_{opt} with the lowest $f_{l_2}(z_{opt})$.

2. Random initial estimate of z, followed by non-linear l_1 optimization of the residuals. We repeat this process k times and choose the solution z_{opt} with the lowest $f_{l_1}(z_{opt})$.

3. A single run of RANSAC algorithm using the minimal solver, followed by non-linear optimization using the measurements that were considered to be inliers by the RANSAC-algorithm.

We evaluate the performance of the algorithm by studying the estimates of the microphone rigs $r = (\mathbf{r}_1, \dots \mathbf{r}_m)$. We measure this error as

$$d(r, (\tilde{r})) = \sqrt{\sum_i |\mathbf{r}_i - \tilde{\mathbf{r}}_i|^2}.$$

If the error is smaller than a threshold (here chosen as 0.05), then the estimate is considered to have succeeded otherwise not.

Outlier Percentage	Random+l_2	Random + l_1	RANSAC + l_2
0	10	17	100
5	0	7	93
10	0	16	87

Table 3: The table shows the success rate in percent of recovering the mic-rig positions for three algorithms for different fraction of outliers i the measurements.

In table 3 the performance for the three algorithm is shown on three datasets of problems, without outliers, with 5 % of outliers and with 10 % outliers respectively. Here, the RANSAC appraoch shows much better results in recovering the mic-rig positions in the presence of outliers. A further bonus is that each measurement gets a classification as either inlier or outlier using the minmial solvers and RANSAC, thus giving a good hypothesis for which measurments are inliers.

A final real example was made using "Spoonphone" ultra-wide-band (UWB) tags and mobile phones as measuring devices, see bottom part of figure 1. The problem is symmetric in terms of tags and phones. Thus we could consider a phone with two UWB units or two such phones mounted on a moving platform moving in an environment with several tags. Alternatively we could consider pairs of UWB tags with known inter-tag distance and a moving phone.

In this experiment we chose the latter setup and placed 3 pairs of tag with known inter-tag distance of 1 meter on the floor. We then measured the distances to these 6 tags at a number of positions using the mobile phone. The measured distances from the 6 tags to a subset of the measurement positions are given in table 4.

By using a minimal solver (3 dual-rig case moving in a plane to 2 points in general position), we obtain an initial estimate of the tags and of the phone motion. The remaining phone-positions are obtained using trilateration. The final estimate, shown in figure 9 is found by non-linear least squares optimization.

6 Conclusion

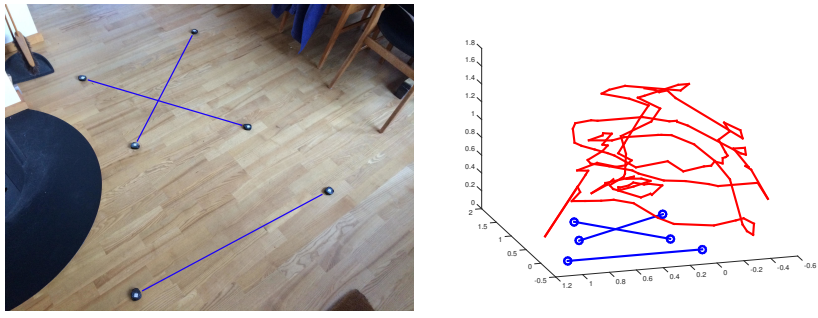
We consider the problems with varying setups as (i) if the internal distances between the microphone nodes are known a priori or not. (ii) if the microphone racks lies in an affine space with different dimension than the sound sources. A

	1	2	3	4	5	6	7
1	2.07	1.43	1.14	0.936	0.669	1.87	0.875
2	2.29	1.96	1.33	1.6	0.911	1.83	0.874
3	1.5	1.56	2.02	2.02	1.83	1.61	1.02
4	1.77	1.25	1.29	1.08	1.02	1.69	0.798
5	1.73	1.6	1.59	-	1.23	1.72	0.782
6	-	1.32	1.98	1.83	1.72	1.72	1.22

Table 4: The table shows the measured distances, using ultra-wide-band techniques, from the 6 tags to a subset of the measurement positions.

set of minimal configurations (i.e. minimal number of microphones and sources) are derived for which the problem is solvable. Such minimal configuration depend on the subspace in which microphone and source lay (3D-3D, 2D-3D, 3D-2D), and on the knowledge of the distance between each microphone couple. For each minimal configuration a non-iterative solving strategy is devised, based on rank constraints of a modified matrix of TOAs, and subsequent polynomial system. Synthetic experiments are carried out for a subset of configurations to assess the reconstruction accuracy vs Gaussian noise on measurements, showing good results. One of the solver is further used on data sets containing outliers in a RANSAC setting, showing good results and the feasibility of using these minimal solvers as a part in finding speaker and mic-rack positions in the presence of severe outliers, a common problem in TOA and TDOA measurements. A real experiment using ultra-wideband shows a reasonable solution. Finally a real sound experiment is carried out for one minimal configuration. Results are good with 5cm RMSE error.

It is difficult to give comparison evaluation, since this work assumes a particular microphone array setup that each pair of microphones are sat on the rack which has same length, but similar setup not found in other literature. It is also hard to compare performance with the case of single microphones distribution if each pair of microphones not distributed at same inter-distance. In our solver we used extra equations comparing to [20] which are the equations (E) or (F). In our settings we have assumption that there is fixed length between each pair of microphones, as if a robot or receiver with a dual-pair of microphones have when moving between measurement positions. If one could solve the above setup for not fixed rack length, i.e. each pair of microphones does not necessarly have



(a) (b)

Figure 9: a) Image showing the position of the 6 ultra-wide-band tags on the floor. The known inter-tag distances are shown as blue lines. b) 3D positions (blue) of the six ultra-wideband tags and the motion (red) of the ultra-wideband equipped phone as estimated from the distance measurements.

same rack length, then it would be possible to compare it with single microphone distribution problem, here we have even number of microphones. This could be the different view of solving calibration problem for considering microphones by pair.

The addressed problem is more specialized than the general TOA problem which was discussed in [20], presenting a closed-form minimum solution. However, we specify the problem further and it reduces the required numbers of signals and receivers even more. We see this problem as a building block for TDOA based self-calibration problem of dual microphone racks, and we believe it can be used to further analyze problem within radio, Wi-Fi and ultrasound.

Acknowledgment

The work was supported by the strategic research projects ELLIIT and eSSENCE and the Swedish Foundation for Strategic Research projects ENGROSS and VINST (grants no. RIT08-0075 and RIT08-0043), and the Swedish Research Council.

References

[1] Magnus Oskarsson, Kalle Åström, and Anna Torstensson, "Prime rigid graphs and multidimensional scaling with missing data," in *International Conference on Pattern Recognition*, 2014.

[2] Carlile Lavor, Leo Liberti, and Nelson Maculan, "Molecular distance geometry problem molecular distance geometry problem," in *Encyclopedia of Optimization*, Christodoulos A. Floudas and Panos M. Pardalos, Eds., pp. 2304–2311. Springer US, 2009.

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

[4] A. Cirillo, R. Parisi, and A. Uncini, "Sound mapping in reverberant rooms by a robust direct method," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 31 2008-april 4 2008, pp. 285 –288.

[5] M. Cobos, A. Marti, and J.J. Lopez, "A modified srp-phat functional for robust real-time sound source localization with scalable spatial sampling," *Signal Processing Letters, IEEE*, vol. 18, no. 1, pp. 71 –74, jan. 2011.

[6] Hoang Do, H.F. Silverman, and Ying Yu, "A real-time srp-phat source location implementation using stochastic region contraction(src) on a large-aperture microphone array," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, april 2007, vol. 1, pp. I–121 –I–124.

[7] Zhayida Simayijiang, Fredrik Andersson, Yubin Kuang, and Kalle Åström, "An automatic system for microphone self-localization using ambient sound," in *European Signal Processing Conference (Eusipco 2014)*, 2014.

[8] Erik Ask, Yubin Kuang, and Kalle Åström, "A unifying approach to minimal problems in collinear and planar tdoa sensor network self-calibration," in *European Signal Processing Conference (EUSIPCO 2014)*, 2014.

[9] G. Young and A.S. Householder, "Discussion of a set of points in terms of their mutual distances," *Psychometrika*, vol. 3, no. 1, pp. 19–22, 1941.

[10] S. T. Birchfield and A. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *IEEE transactions on Speech and Audio Processing*, vol. 13, no. 5, 2005.

[11] D. Niculescu and B. Nath, "Ad hoc positioning system (aps)," in *GLOBECOM-01*, 2001.

[12] E. Elnahrawy, Xl. Li, and R. Martin, "The limits of localization using signal strength," in *SECON-04*, 2004.

[13] 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*, vol. 13, no. 1, 2005.

[14] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan*, March 2012.

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

[16] R. Biswas and S. Thrun, "A passive approach to sensor network localization," in *IROS 2004*, 2004.

[17] J. Wendeberg, F. Hoflinger, C. Schindelhauer, and L. Reindl, "Anchor-free tdoa self-localization," in *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on*, Sept. 2011, pp. 1 –10.

[18] P. Biswas, T.C. Lian, T.C. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Transactions on Sensor Networks (TOSN)*, vol. 2, no. 2, pp. 188–220, 2006.

[19] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, April 2005.

[20] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Processing*, 2013.

[21] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[22] Yubin Kuang and Kalle Åström, "Stratified sensor network self-calibration from tdoa measurements," in *21st European Signal Processing Conference*, 2013.

[23] Yubin Kuang, Erik Ask, Simon Burgess, and Karl Åström, "Understanding toa and tdoa network calibration using far field approximation as initial estimate," in *ICPRAM*. 2012, pp. 590–596, SciTePress.

[24] Simon Burgess, Yubin Kuang, Johannes Wendeberg, Kalle Åström, and Christian Schinelhauer, "Minimal solvers for unsynchronized tdoa sensor network calibration," in *9th International Symposium on Algorithms and Experiments for Sensor Systems, Wireless Networks and Distributed Robotics*, 2013.

[25] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[26] Fredrik Gustafsson, *Statistical sensor fusion*, Studentlitteratur, Lund, 2010.

[27] N. D. Gaubitch, W. B. Kleijn, and R. Heusdens, "Auto-localization in ad-hoc microphone arrays," in *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Processing*, 2013.

[28] Martin Byröd, Klas Josephson, and Kalle Åström, "Fast and stable polynomial equation solving and its application to computer vision," *Int. Journal of Computer Vision*, vol. 84, no. 3, pp. 237–255, 2009.

[29] Arun K.S., Huang T.S., and Blostein S.D., "Least-squares fitting of two 3-d point sets," *IEEE Transactions on pattern analysis and machine intelligence*, 1987.

E

Paper E

A Complete Characterization and Solution to the Microphone Position Self-Calibration Problem

Yubin Kuang, Simon Burgess, Anna Torstensson, and Kalle Åström

Centre for Mathematical Sciences, Lund University, Lund, Sweden

Abstract

This paper presents a complete characterization and solution to microphone position self-calibration problem for time-of-arrival (TOA) measurements. This is the problem of determining the positions of receivers and transmitters given all receiver-transmitter distances. Such calibration problems arise in applications such as calibration of radio antenna networks, audio or ultra-sound arrays and WiFi transmitter arrays. We show for what cases such calibration problems are well-defined and derive efficient and numerically stable algorithms for the minimal TOA based self-calibration problems. The proposed algorithms are non-iterative and require no assumptions on the sensor positions. Experiments on synthetic data show that the minimal solvers are numerically stable and perform well on noisy data. The solvers are also tested on two real datasets with good results.

Key words: Time-of-Arrival, Network Calibration, Microphone Self-Calibration, Minimal Solver.

1 Introduction

The problem of sensor network self-calibration is essential to localization and navigation. In this paper we focus on the time-of-arrival (TOA) based self-calibration problem, i.e. the problem of determining the positions of a number of receivers and transmitters given all receiver-transmitter distances. This problem has certain similarities to the problem of determining a set of points given all inter-point distances, which is usually solved using multi-dimensional scaling [1]. Such problems are of general interest in visualization and analysis of large datasets and for many geometric problems. It also relates to the study of sensor networks under rigid graph theory [2, 3] where general graph structure is of interest. The TOA-based self-calibration problem studied here corresponds to a special case - bipartite graph [4]. It is important for network calibration using e.g. microphone arrays, given recordings of sounds emitted at unknown locations, to microphones at unknown positions, determine both sound emission positions and microphone locations.

Such problems could be solved using alternative techniques, such as manually measuring all distances between microphones or using computer vision. Examples of such approaches are given in [5–10]. Here we argue that efficient solution to the TOA-based self-calibration problem opens up new technological possibilities e.g. calibration of a sensor network on the fly, determining of reflections of receivers and transmitters while moving in an unknown terrain etc. The solution of the problem is also of great theoretical interest and the solution techniques are interesting per se.

Iterative methods exist for TOA or time-difference-of-arrival (TDOA) based self-calibration [11, 12]. However, such methods can get stuck in local minima which are dependent on initialization. Non-iterative methods for initializing network self-calibration have received less attention. For general graph structure, one can relax the TOA-based calibration problem as a semi-definite program [13]. For bipartite graph, initialization of TOA-based bipartite networks studied in [14], where solutions to the minimal case of 3 transmitters and 3 receivers in the plane is given. Initialization of time-difference-of-arrival (TDOA) networks is studied in [15] where a solution to non-minimal case of 10 receivers and 4 transmitters in 3D for TOA problem was also derived. In [16, 17] a solution is given to the TOA based self-calibration problem, if one may additionally assume that one of the receivers coincide with the position of one of the transmitters. In [18] and refined in [19] a far field approximation was utilized to initialize both TOA and

TDOA problems. In [20, 21], algorithms for far-field unsynchronized receivers were also proposed. All these previous works attempt to solve the problem with either minimal or close to minimal data. Studying these minimal cases is both of theoretical importance and essential to develop fast stable algorithms suitable in RANSAC [22] and other solution schemes.

In this paper, we completely characterize the TOA based self-calibration problem in three dimensions. It is shown that such problems are well-defined for m receivers and n transmitters if and only if $m \geq 4$, $n \geq 4$, $m+n \geq 10$. We present efficient, numerically stable and non-iterative algorithms for such problems. In particular, we study the minimal problem of $(m = 6, n = 4)$ (or $(m = 4, n = 6)$ which is identical because of symmetry). We show that this problem has in general 38 solutions and present an algorithm for determining these 38 solutions given an arbitrary 4×6 matrix of distance measurements. Furthermore we study the the problem of $(m = 5, n = 5)$. We show that each 5×5 matrix must fulfill one constraint. But as long as this one constraint is fulfilled the problem is minimal and has 42 solutions. Also for this problem we provide a fast and numerically efficient algorithm. To the best our knowledge, our algorithms are the first to give practical numerical solution for minimal TOA calibration problems in 3D. We also extend the solution scheme to overdetermined cases and sketch how the technique could be extended to other dimensions.

2 The TOA-based Calibration Problem

Let $\mathbf{r}_i$, $i = 1, \ldots, m$ and $\mathbf{s}_j$, $j = 1, \ldots, n$ be the spatial coordinates of m receivers (e.g. microphones) and n transmitters (e.g. sound events), respectively. For measured time of arrival t_{ij} from transmitter $\mathbf{r}_i$ and receiver $\mathbf{s}_j$, we have $vt_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2$ where v is the speed of measured signals. We assume that v is known and constant, and that we at each receiver can distinguish which transmitter j each event is originating from. This can be done e.g. if the signals are temporally separated or by using different frequencies. We will in the sequel work with the distance measurements $d_{ij} = vt_{ij}$. The TOA calibration problem can then be defined as follows.

Problem 2.1. (Time-of-Arrival Self-Calibration) Given absolute distance measurements d_{ij} determine receiver positions $\mathbf{r}_i$, $i = 1, \ldots, m$ and transmitter positions $\mathbf{s}_j, j = 1, \ldots, n$ such that $d_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2$.

Note that for such problems, one can only reconstruct locations of receivers and transmitters up to euclidean transformation and mirroring, henceforth referred to as the gauge freedom. In the following discussion, we assume that the dimensionality of the affine space spanned by $\mathbf{r}_i$ and $\mathbf{s}_j$ is the same and it is denoted as K, typical values in practice is $K = 3$ for transmitters and receivers in general 3D positions. The minimal configurations for different dimensions have previously been determined in [4, 14]. It is relatively straightforward to calculate the number of degrees of freedom in the measurements, mn and the number of degrees of freedom in the manifold of unknown parameters, e.g. $3n + 3m - 6$ for TOA in 3D. To be more precise one has to study the equations using algebraic geometry, to make certain that there are no anomalies in the set of equations.

2.1 Analysis of the TOA-based Calibration Problem

In this section, we analyze the minimal cases for TOA-based self-calibration problem. Here we define a *minimal case* to a problem as the case that consists of the minimal set of constraints or equations such that the problem generally has finite and > 0 number of solutions. We start by deriving new sets of equations.

Since the distance measurements are assumed to be real and positive one does not lose any information by squaring them, i.e.

$$d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j) = \mathbf{r}_i^T\mathbf{r}_i + \mathbf{s}_j^T\mathbf{s}_j - 2\mathbf{r}_i^T\mathbf{s}_j. \tag{1}$$

Notice that these are now polynomial equations in the unknowns. The problem is significantly easier to analyze and solve by forming new equations according to the following linear combinations of d_{ij}^2:

$$\begin{pmatrix} d_{11}^2 & d_{12}^2 - d_{11}^2 & \dots & d_{1n}^2 - d_{11}^2 \\ d_{21}^2 - d_{11}^2 & & & \\ \dots & & \tilde{\mathbf{D}} & \\ d_{m2}^2 - d_{11}^2 & & & \end{pmatrix}, \tag{2}$$

where $\tilde{\mathbf{D}}$ is a $(m-1)(n-1)$ matrix with entries as $\tilde{d}_{ij} = d_{i,j}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2$, with $i = 2, \dots m$ and $j = 2, \dots, n$.

These new mn equations are equivalent to the mn equations formed by d_{ij}^2. The new ones are in fact an invertible linear combinations of the old ones.

These equations are of four types:

(A) 1 equation $d_{11}^2 = (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1)$.

(B) $n-1$ equations of type $d_{1j}^2 - d_{11}^2 = \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 - 2\mathbf{r}_1^T(\mathbf{s}_j - \mathbf{s}_1)$, for $j = 2, \ldots n$.

(C) $m-1$ equations of type $d_{i1}^2 - d_{11}^2 = \mathbf{r}_i^T\mathbf{r}_i - \mathbf{r}_1^T\mathbf{r}_1 - 2(\mathbf{r}_i - \mathbf{r}_1)^T\mathbf{s}_1$, for $i = 2, \ldots m$.

(D) $(m-1)(n-1)$ equations of type $d_{ij}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2 = -2(\mathbf{r}_i - \mathbf{r}_1)^T(\mathbf{s}_j - \mathbf{s}_1)$, for $i = 2, \ldots m, j = 2, \ldots n$.

Without loss of generality we may assume that $m \geq n$. It turns out that the characterization of the problem depends on the affine span of the transmitters and the receivers. For simplicity we will in the following concentrate on 3D problems and will assume that the affine span of both the transmitters and of the receivers are of dimension 3. Notice, however, that much of the argumentation and theory is straightforward to generalize to other dimensions. A brief sketch on what can be said in other dimensions is given in Section 2.4.

The solution strategy is to use the equations of type D first and use factorization techniques to solve for **r**'s and **s**'s up to a translation vector **b** (3 degrees of freedom) and an affine deformation **L** (6 degrees of freedom up to an unknown rotation). By a clever choice of parametrization of the problem it can be shown that the equations of Type C are linear in the unknowns and the equations of Types A and B can be used to form polynomial equations. We will then show that such problems are well defined and can be solved if $m \geq 4$, $n \geq 4$ and $m + n \geq 10$. The interesting minimal cases are thus $m = 6, n = 4$ (and $m = 4$, $n = 6$) as well as $m = 5$, $n = 5$. It was shown in [4] that these corresponding to rigid cases for bipartite graphs in 3D.

The factorization step can be understood as follows. Let $\mathbf{R}_i = \left[(\mathbf{r}_i - \mathbf{r}_1)\right]$ and $\mathbf{S}_j = \left[-2(\mathbf{s}_j - \mathbf{s}_1)\right]$. The equations of type D can be written as $\tilde{\mathbf{D}} = \mathbf{R}^T\mathbf{S}$ with $\mathbf{R}_i$ as columns of $\mathbf{R}$ and $\mathbf{S}_j$ as columns of $\mathbf{S}$. The ranks of $\mathbf{R}$ and $\mathbf{S}$ depends on the dimensionality of the affine span of the receivers and the transmitters respectively. As we assume that both of these are 3, then the matrix $\tilde{\mathbf{D}}$ also has rank 3. This also implies that in order to solve the problem, it is required that $m \geq 4$ and $n \geq 4$. By factorizing $\tilde{\mathbf{D}}$ which is of rank 3 using e.g. singular value decomposition, we can compute the vectors to all receivers and transmitters from unknown initial/reference positions ($\mathbf{r}_1$ and $\mathbf{s}_1$) up to an unknown full-rank 3×3 transformation $\mathbf{L}$ such that $\tilde{\mathbf{D}} = \tilde{\mathbf{R}}^T\mathbf{L}^{-1}\mathbf{L}\tilde{\mathbf{S}} = \mathbf{R}^T\mathbf{S}$.

To solve for the unknown transformation and reference positions, we now utilize the nonlinear constraints in equations of Type A, B and C. First we can

fix the translational part of the gauge freedom by choosing the location $\mathbf{r}_1$ at the origin. Given that $\mathbf{R} = \mathbf{L}^{-T}\tilde{\mathbf{R}}$ and $\mathbf{S} = \mathbf{L}\tilde{\mathbf{S}}$, we can parameterize $\mathbf{s}_1$ as $\mathbf{Lb}$ where $\mathbf{b}$ is a 3×1 vector. This gives

$$\begin{aligned}
&\mathbf{r}_1 = \mathbf{0},\ \mathbf{s}_1 = \mathbf{Lb} \\
&\mathbf{r}_i = \mathbf{L}^{-T}\tilde{\mathbf{R}}_i,\ i = 2 \dots m \\
&\mathbf{s}_j = \mathbf{L}(\tilde{\mathbf{S}}_j^* + \mathbf{b}),\ j = 2 \dots n,
\end{aligned} \tag{3}$$

where $\tilde{\mathbf{S}}_j^* = \tilde{\mathbf{S}}_j/(-2)$. Using this parametrization the equations of type (A), (B) and (C) become

$$\begin{aligned}
d_{11}^2 &= (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1) = \mathbf{s}_1^T\mathbf{s}_1 \\
&= \mathbf{b}^T\mathbf{L}^T\mathbf{Lb}, && (4) \\
d_{1j}^2 - d_{11}^2 &= \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 \\
&= \tilde{\mathbf{S}}_j^{*T}\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j^* + 2\mathbf{b}^T\mathbf{L}^T\mathbf{L}\tilde{\mathbf{S}}_j^*, && (5) \\
d_{i1}^2 - d_{11}^2 &= \mathbf{r}_i^T\mathbf{r}_i - 2\mathbf{r}_i^T\mathbf{s}_1 \\
&= \tilde{\mathbf{R}}_i^T(\mathbf{L}^T\mathbf{L})^{-1}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i. && (6)
\end{aligned}$$

Observe that all the constraints involve only $\mathbf{L}^T\mathbf{L}$ (and its inverse) and $\mathbf{b}$. By representing $(\mathbf{L}^T\mathbf{L})^{-1}$ with a symmetric matrix $\mathbf{H}$ parameterized with 6 unknowns, the constraints in (4), (5) and (6) can then be simplified as

$$\begin{aligned}
d_{11}^2 &= \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}, && (7) \\
d_{1j}^2 - d_{11}^2 &= \tilde{\mathbf{S}}_j^{*T}\mathbf{H}^{-1}\tilde{\mathbf{S}}_j^* + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{S}}_j^*, && (8) \\
d_{i1}^2 - d_{11}^2 &= \tilde{\mathbf{R}}_i^T\mathbf{H}\tilde{\mathbf{R}}_i - 2\mathbf{b}^T\tilde{\mathbf{R}}_i. && (9)
\end{aligned}$$

With this parameterization, there are in total 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively). By utilizing $\mathbf{H}^{-1} = \text{adj}(\mathbf{H})/\det(\mathbf{H})$, where $\text{adj}(\mathbf{H})$ is the adjoint of $\mathbf{H}$, we can multiply equations in (7) and (8) by $\det(\mathbf{H})$ to rewrite them as polynomials equations. In this case, we have $(n + m - 1)$ equations, among which the $(m - 1)$ equations in (9) are linear, the $(n - 1)$ equations in (8) are polynomial equations of degree 3 and Equation (7) is of degree 4. Thus we need $n + m - 1 \geq 9$ or $n + m \geq 10$ in order to solve for the 9 unknowns. Since both $m \geq 4$ and $n \geq 4$ there are two minimal cases $6r/4s$ ($4r/6s$) and $5r/5s$.

2.2 Solving the Polynomial System

For the minimal case of 6 receivers and 4 transmitters, there are 5 linear equations of type C. By linear elimination we can express $\mathbf{H}$ and $\mathbf{b}$ in terms of $9 - 5 = 4$ unknowns $\mathbf{x} = (x_1, x_2, x_3, x_4)$. We now obtain four equations

$$\det(\mathbf{H})d_{11}^2 = \mathbf{b}^T \operatorname{adj}(\mathbf{H})\mathbf{b} \tag{10}$$

$$\det(\mathbf{H})(d_{12}^2 - d_{11}^2) = \tilde{\mathbf{S}}_2^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_2^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_2^* \tag{11}$$

$$\det(\mathbf{H})(d_{13}^2 - d_{11}^2) = \tilde{\mathbf{S}}_3^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_3^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_3^* \tag{12}$$

$$\det(\mathbf{H})(d_{14}^2 - d_{11}^2) = \tilde{\mathbf{S}}_4^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_4^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_4^* \tag{13}$$

in the four unknowns. Here both $\mathbf{H}$ and $\mathbf{b}$ depend on $\mathbf{x}$. Using tools from algebraic geometry it can be shown that the solution set to equations (9-12) in general consists of a set of dimension 1 (a curve) of 'false' solutions that fulfill $\det(H) = 0$ and 38 points. This is done by running the system of equations in Macaulay2 [23] over the field of $\mathbb{Z}^p$, where p is a large prime number and with coefficients initialized randomly. To remove the one-dimensional curve of false solutions we employ a saturation technique as follows. We rewrite the equations using an additional unknown z and an additional equation $\det(\mathbf{H}) = z$, i.e.

$$zd_{11}^2 = \mathbf{b}^T \operatorname{adj}(\mathbf{H})\mathbf{b} \tag{14}$$

$$z(d_{12}^2 - d_{11}^2) = \tilde{\mathbf{S}}_2^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_2^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_2^* \tag{15}$$

$$z(d_{13}^2 - d_{11}^2) = \tilde{\mathbf{S}}_3^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_3^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_3^* \tag{16}$$

$$z(d_{14}^2 - d_{11}^2) = \tilde{\mathbf{S}}_4^{*T} \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_4^* + 2\mathbf{b}^T \operatorname{adj}(\mathbf{H})\tilde{\mathbf{S}}_4^* \tag{17}$$

$$\det(\mathbf{H}) = z \tag{18}$$

We then multiply all equations with monomials in $\mathbf{x}$ up to degree 3 and keep the highest degree of z as 1. By doing this one can construct 315 equations involving 330 monomials which do not contain z and 70 monomials that do contain z. These equations can be represented by a sparse coefficient matrix $\mathbf{M} = [\mathbf{M}_0 \ \ \mathbf{M}_z]$ of size 315×400, where the coefficients corresponding to monomials without z are in $\mathbf{M}_0$ and those corresponding to monomials with z are in $\mathbf{M}_z$. After multiplication with $\mathbf{Q}^T$, where $\mathbf{QR} = \mathbf{M}_0$ is the QR-factorization of $\mathbf{M}_0$, we obtain

$$\mathbf{Q}^T\mathbf{M} = [\mathbf{R} \ \ \mathbf{Q}^T\mathbf{M}_z].$$

Here the last 51 rows of $\mathbf{R}$ is zero and thus the last 51 equations can all be written $zf_k(\mathbf{x}) = 0$. After division with z, we obtain 51 equations of degree 4 in $\mathbf{x}$. It can be shown that this solution set to these equations consist of 38 points.

We then use these equations for solving for the 38 solutions using a technique described in [24]. This again involves multiplying the 51 equations with monomials, generating corresponding coefficient matrix, row manipulation in order to generate a 38×38 matrix, whose eigenvalues and eigenvectors contain the solution to the system of polynomial equations. For each such solution, we then calculate $\mathbf{H}$ and $\mathbf{b}$ and then generate the solutions for $\mathbf{r}_i$ and $\mathbf{s}_j$ according to (3), finding $\mathbf{L}^{-1}$ by e.g. cholesky factorization of $\mathbf{H}$. $\mathbf{L}$ is thus only determined up to a matrix R where $R^T R = I$, which coincides with the gauge freedom of rotating and/or mirroring our solution. The 4r/6s case can be solved in the same way by first transposing the measurement matrix.

The case of 5 receivers and 5 transmitters is interesting. It is an overdetermined case in the sense that there are 25 measurements and 24 degrees of freedom in the solutions set. There is thus one constraint that has to be satisfied, i.e. the constraint that the 4×4 matrix $\tilde{\mathbf{D}}$ has determinant zero. However for all such data, the problem of determining $\mathbf{H}$ and $\mathbf{b}$ is minimal. There are $m + n - 1 = 9$ equations (1 of Type A, 4 of Type B and 4 of Type C) and 9 unknowns. We follow a similar solution scheme as for the (6r/4s) case. By linear elimination using the 4 linear constraints of type C, we can express $\mathbf{H}$ and $\mathbf{b}$ in terms of $9 - 4 = 5$ unknowns $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$. The remaining five constraints (1 of Type A, and 4 of Type B) give a polynomial system with 42 solutions after a saturation procedure similar to the previous case. Again we use the action matrix approach to produce a numerically stable and efficient solution scheme.

2.3 Overdetermined cases

For overdetermined cases ($m \geq 4, n \geq 4, m + n > 10$), the solver can be based on solving a minimal case, extending with trilateration, followed by non-linear optimization to obtain the maximum likelihood estimate. An alternative is to find the best rank 3 approximation of $\tilde{\mathbf{D}}$, solution of $\mathbf{H}$ and $\mathbf{b}$ using algebraic methods, and then again followed by non-linear optimization. An advantage with the former approach is that it can more easily be modified using RANSAC to remove potential outliers in the measurement matrix $\mathbf{D}$.

2.4 Higher and lower dimensional cases

The ideas presented here can relatively easy to generalized to other dimensions. The one-dimensional case is trivial. Only one measurement is needed to solve

for the problem. In two dimensions the same approach can be used to show that one needs $m \geq 3, n \geq 3, m+n \geq 6$, which indicate only one minimal problem $3s/3r$. This was in fact solved with a different approach by Stewenius and Nister in [14] . The problem has in general 4 solutions. For dimension 4, the analysis gives $m \geq 5, n \geq 5, m+n \geq 15$, which gives minimal cases $5s/10r, 6s/9r, 7s/8r$,

3 Experiments

We test our proposed algorithms on both synthetic and real data. The recovered sensor positions are up to unknown rotation and translation. We first determine the rotation and translation with least square fitting over the ground truth positions and then compute the errors. For synthetic data, we simulate the positions of receivers and transmitters as 3D points with independent Gaussian distribution of zero mean and identity covariance matrix. The 6r/4s solver always produce 38 solutions although some of these could be complex. Similarily the 5r/5s always produce 42 solutions. There are two steps in the algorithm where the solution could become non-real. First a solution to the system of polynomial equation could be non-real and secondly even if this solution is real the matrix $\mathbf{H}$ although real, could be indefinite in which case the Cholesky factorization becomes non-real. In Fig.1 (Left) is shown a histogram over 5000 simulations. As can be seen in the figure, for both solvers, there are usually between 14 and 24 real solutions to the system of polynomial equations, whereas only a few (most often less than 6) of these produce positive definite matrices $\mathbf{H}$), so that there is typically less than 6 real and thus valid solutions. This number is however data dependent. For noise-free synthetic data, we can see in Fig.1 (Right) that both the 6r/4s solver and 5r/5s solver are numerically stable. We also tested both solvers on data under different noise level and we observe that the solvers gives fairly good solutions under reasonable level of noise (Fig.2, left, solid lines). Using the solutions from the minimal solvers as initial solutions, we also apply nonlinear optimization step where we minimize $\sum_{ij} \left(d_{ij} - (||\mathbf{r}_i - \mathbf{s}_j||_2)\right)^2$ (Fig.2, dash lines). We also test the solvers on over-determined cases with fixed noise level, $m = 10$ and varying n (Fig.2, right). We can see that as n increases, both solvers gives better initial solutions for the reconstruction. The current implementation for both solvers run at around 1 second on a Macbook Air (1.8 GHz Intel Core i5 and 8 GB memory). The solvers are available for download at `http://www2.maths.lth.se/vision/downloads/`.

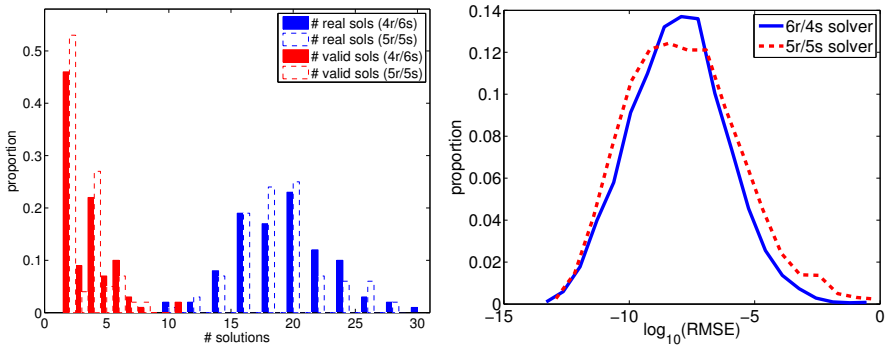


Figure 1: Minimal solver performance on 5000 noise-free random synthetic TOA problems. Left: Distribution of the number of real and valid solutions each run produces, showing the relative frequency of number of real and valid solutions among the 38 (or 42) solutions. Right: the error distribution (RMSE) of reconstructed positions of microphone and sound sources.

For real experiments, we have used a publicly available dataset [16] for comparison as well as our own TOA measurements. In the dataset [16], the distances between the 8 microphones and 21 sounds are estimated based on the time-of-arrival measurements. The first microphone is assumed to be at the same location as first sound. For our algorithm, no such assumption is needed. To verify this, we simply remove the distance measurements corresponding to the first microphone and first sound, which gives us a 7×20 matrix. For this reduced set of measurements, the root mean square errors (RMSE) of our reconstructed positions of microphones and sound sources after non-linear iterative optimization are $0.0083m$ and $0.0108m$, respectively. This is similar to the accuracy [16] that achieves ($0.0091m$ for microphones and $0.0111m$ for sound sources) with the additional assumption. For the full set of data (8 microphones and 21 sounds), our solvers also gives similar errors as in [16]. Another set of real data was obtained using seven T-bone MM-1 microphones and five Roxcore portable speakers, connected to a Fast Track Ultra 8R sound card in an indoor environment, with speakers and microphones placed in an approximate $1.5 \times 1.5 \times 1.5$ m^3 volume (Fig.3, Left). TOA measurements were obtained by heuristically matching sounds from different speakers to sound flanks recorded from different microphones. For this set of measurement, we also reconstructed the scene using computer vision based

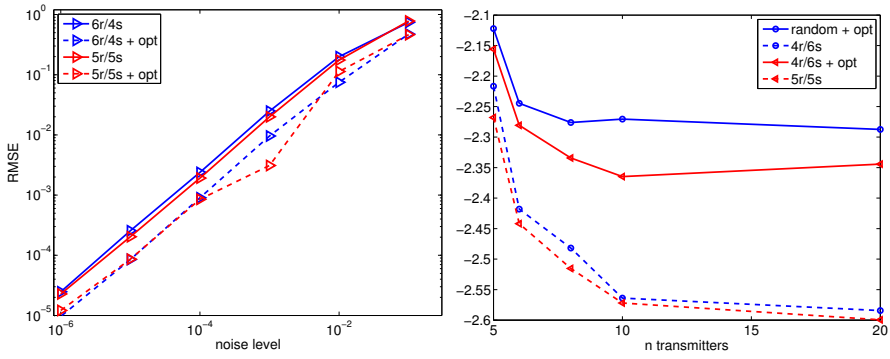


Figure 2: Performance on noisy synthetic data - average errors (RMSE) of reconstructed positions of receivers and transmitters Left: minimal solvers (4r/6s and 5r/5s) under varying noise levels. Right: 10 receivers and varying number of transmitters under Gaussian noise with standard deviation 2×10^{-3}.

algorithms as ground truth. The reconstruction (Fig.3, Right) when compared to the vision-based reconstruction has RMSE $0.0088m$ and $0.0131m$ for microphones and speakers respectively.

4 Relation to Prior Work

The most related works to ours are [16, 17], where the same factorization step is described. However, instead of solving the original problem, they solve a problem with an additional assumption that one receiver and one transmitter have identical positions. This work does not need this constraint, and thus can solve a wider set of problems. Close to our setting are also [14], where the minimal TOA calibration problem in 2D is solved, using three receivers and transmitters each, [15] which solves as a by-product the TOA calibration problem in 3D with a non-minimal configuration of 10 receivers and 4 transmitters and [19] which solves the TOA calibration problem assuming a far field approximation with a minimal configuration of 3 receivers and 6 transmitters. Thus no previous work has solved the minimal cases for general TOA-based self-calibration in 3D.

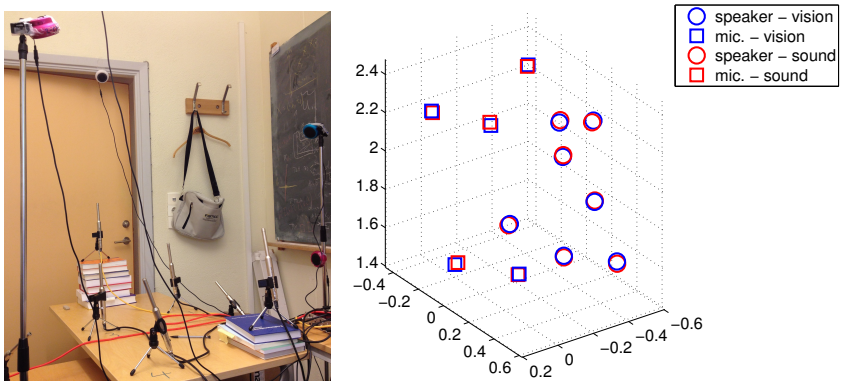


Figure 3: Real microphone and speaker calibration setup, Left: the setup of microphones and speakers in an office, Right: the reconstructed sensor positions using TOA measurements (red) aligned with the positions estimated based on computer vision (blue).

5 Conclusions

In this paper, we completely characterize the TOA base self-calibration problem in three dimensions. It is shown that such problems are well-defined for m receivers and n transmitters if and only if $m \geq 4$, $m \geq 4$, $m+n \geq 10$. We present practical and non-iterative solution algorithms for such problems. In particular, we present algorithms for solving the $(m = 6, n = 4)$ case (38 solutions) and the $(m = 5, n = 5)$ (42 solutions). For overdetermined cases we present two alternative approaches. The solution technique is general and can in principle be applied to higher dimensional problems. In the paper we show the applicability of the techniques to both simulated and real data. In the future it would be interesting to study how such algorithms could be used to further analyze problems within radio, Wi-Fi and ultrasound.

Ackgnowledgements

The research leading to these results has received funding from the strategic research projects ELLIIT and eSSENCE, and Swedish Foundation for Strategic Research projects ENGROSS and VINST(grant no. RIT08-0043).

References

[1] G. Young and A.S. Householder, "Discussion of a set of points in terms of their mutual distances," *Psychometrika*, vol. 3, no. 1, pp. 19–22, 1941.

[2] Leonard Asimow and Ben Roth, "The rigidity of graphs, ii," *Journal of Mathematical Analysis and Applications*, vol. 68, no. 1, pp. 171–190, 1979.

[3] T. Eren, OK Goldenberg, W. Whiteley, Y.R. Yang, A.S. Morse, BDO Anderson, and PN Belhumeur, "Rigidity, computation, and randomization in network localization," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*. IEEE, 2004, vol. 4, pp. 2673–2684.

[4] E.D. Bolker and B. Roth, "When is a bipartite graph a rigid framework," *Pacific J. Math*, vol. 90, no. 1, pp. 27–44, 1980.

[5] S. T. Birchfield and A.. Subramanya, "Microphone array position calibration by basis-point classical multidimensional scaling," *IEEE transactions on Speech and Audio Processing*, vol. 13, no. 5, 2005.

[6] D. Niculescu and B. Nath, "Ad hoc positioning system (aps)," in *GLOBECOM-01*, 2001.

[7] E. Elnahrawy, Xl. Li, and R. Martin, "The limits of localization using signal strength," in *SECON-04*, 2004.

[8] 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*, vol. 13, no. 1, 2005.

[9] J. Sallai, G. Balogh, M. Maroti, and A. Ledeczi, "Acoustic ranging in resource-constrained sensor networks," in *eCOTS-04*, 2004.

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

[11] N.B. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller, "Anchor-free distributed localization in sensor networks," in *Proceedings of the 1st international conference on Embedded networked sensor systems*. ACM, 2003, pp. 340–341.

[12] R. Biswas and S. Thrun, "A passive approach to sensor network localization," in *IROS 2004*, 2004.

[13] P. Biswas, T.C. Lian, T.C. Wang, and Y. Ye, "Semidefinite programming based algorithms for sensor network localization," *ACM Transactions on Sensor Networks (TOSN)*, vol. 2, no. 2, pp. 188–220, 2006.

[14] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, APR 2005.

[15] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[16] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, "A closed form solution to the microphone position self-calibration problem," in *37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan*, March 2012.

[17] Marco Crocco, Alessio Del Bue, and Vittorio Murino, "A bilinear approach to the position self-calibration of multiple sensors," *Trans. Sig. Proc.*, vol. 60, no. 2, pp. 660–673, feb 2012.

[18] S. Thrun, "Affine structure from sound," in *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, Cambridge, MA, 2005, MIT Press.

[19] Yubin Kuang, Erik Ask, Simon Burgess, and Karl Åström, "Understanding TOA and TDOA Network Calibration using Far Field Approximation as Initial Estimate," in *ICPRAM 2012 - Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods, Volume 2*. 2012, pp. 590–596, SciTePress.

[20] J. Wendeberg, F. Hoflinger, C. Schindelhauer, and L. Reindl, "Anchor-free tdoa self-localization," in *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on*, sept. 2011, pp. 1 –10.

[21] Simon Burgess, Yubin Kuang, and Karl Åström, "Node localization in unsynchronized time of arrival sensor networks," in *21st International Conference on Pattern Recognition (ICPR 2012), Proceedings of*. 2012, pp. 2042–2046, International Association for Pattern Recognition (IAPR) & IEEE.

[22] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–95, 1981.

[23] D. Grayson and M. Stillman, "Macaulay 2," Available at http://www.math.uiuc.edu/Macaulay2/, 1993-2002, An open source computer algebra software.

[24] Martin Byröd, Klas Josephson, and Kalle Åström, "Fast and stable polynomial equation solving and its application to computer vision," *Int. Journal of Computer Vision*, vol. 84, no. 3, pp. 237–255, 2009.

[25] Simon Burgess, Yubin Kuang, and Karl Åström, "Pose estimation from minimal dual receiver configurations," in *21st International Conference on Pattern Recognition (ICPR 2012), Proceedings of*. 2012, pp. 2553–2556, International Association for Pattern Recognition (IAPR) & IEEE.

[26] Erik Ask, Yubin Kuang, and Karl Åström, "Exploiting p-Fold Symmetries for Faster Polynomial Equation Solving," in *21st International Conference on Pattern Recognition (ICPR 2012), Proceedings of*. 2012, pp. 3232–3235, International Association for Pattern Recognition (IAPR) & IEEE.

[27] Yubin Kuang and Kalle Åström, "Numerically stable optimization of polynomial solvers for minimal problems," in *Lecture Notes in Computer Science*, Andrew Fitzgibbon, Ed. 2012, vol. 7574, pp. 100–113, Springerm, Heidelberg.

F

Paper F

Smartphone Positioning in Multi-Floor Environments Without Calibration or Added Infrastructure

Simon Burgess[1], Kalle Åström[1], Mikael Högström[2], Björn Lindquist[2], and Rasmus Ljungberg[2]

[1] *Centre for Mathematical Sciences, Lund University, Lund, Sweden*
[2] *Combain Positioning Solutions, Lund, Sweden*

Abstract

Indoor positioning for smartphone users has received a lot of attention in recent years. While many solutions have been developed, most rely on a need for pre-deployment of infrastructure or collecting ground truth data to train on. In this paper we see what can be done using existing WiFi-infrastructure and Received Signal Strength from these to smartphones, not using any calibration of the signal environment or manually set WiFi positions. We expand on previous work by using a multi-floor model taking into account dampening between floors, and optimize a target function consisting of least squares residuals, to find positions for WiFis and the smartphone measurement locations simultaneously. Pressure sensors are used to do floor estimation. The method was tested inside two multi-story buildings, with 5 stories each, with median errors of smartphone positions of 12.5m and 16.4m and with WiFi median position errors of 7.16m and 19.4m respectively. Correct floor detection was achieved for 96% of all smartphone positions.

1 Introduction

Positioning and navigation has throughout the ages been a key instinct for survival in the natural world. Today, the need for navigation is still an integral part of modern life, present everywhere from finding your keys, to first responders being able to locate where an emergency call was made from with accuracy. The solution of how to do localization and/or mapping depends on algorithms, hardware, environment and the desired precision. A general solution or framework seems to be far away. Outdoors, the Global Navigation Satellite System (GNSS) works well for many applications. Where GNSS signals are compromised, like in close proximity to many tall buildings or indoors, the GNSS generally does not perform well or at all. To alleviate this problem, indoor positioning for users of smart devices such as computers, tablets or phones has attracted a lot of attention in the past few years.

We present a method for simultaneously mapping the signal environment and positioning one or several users from offline data, in multi-story environments. Data used are Received Signal Strength Index (RSSI) from nearby WiFi and Bluetooth devices, a now ubiquitous part of many indoor environments, where GNSS is most commonly not available, or available with limited accuracy. For floor detection, data from pressure sensors is used. The users do not need to provide additional information where they are, but only walk around in the considered environment. No prior calibration of the access points positions or parameters is needed. This approach lends itself well for crowdsourcing future data. When the simultaneous localization of the smart devices and mapping of the signal environment has been done by the methods we present, real-time localization using a trilateration-like approach can readily be done, cf. [1, 2].

Previous approaches for RSSI localization most commonly use a supervised learning approach, where a setup phase of users manually providing location data [2–5], manually installing beacons with known parameters and locations, using floor maps or dead reckoning, cf. [6].

Some previous papers use an unsupervised approach for 2D RSSI localizatin, i.e. finding how a user has moved without prior calibration and mapping the radio environment, but they rely on inertial measurement units such as accelerometer, gyro and magnetometer as well, and/or a annotated floor plan [7–11]. This generally puts heavy constraints on the user holding their phones in a predetermined controlled position, and that sampling frequency should be low and consistent. Another approach for self-calibration is [12], but needs a fair amount

of the access points to be manually calibrated. Other require specialized hardware deployed, [13, 14].

An approach similar to ours is [15], but whereas they use hand-tailored, simulated, GPS signals to fixate the signal environment map, we use only the sparse GPS signals that can be found in or just outside the location. Furthermore, they do not consider the multi-story scenario. As we have been unable to obtain the data sets used in [15], we furthermore provide several data sets with real sparse GPS signals for future benchmarking [1]. We also present models and algorithms for positioning in multi-floor environments, a highly relevant topic for indoor localization considering the 2015 U.S. Federal Communications Commission rules for mobile carriers, [16], which states that an accuracy of 50m horizontally and 3m vertically in 67% of the cases is needed. As outlined in our poster abstract, [17], floor detection of smartphone measurement positions using only the RSSI and sparse GPS values seems to need a supervised or semi-supervised approach. To alleviate this, we use atmospheric pressure sensors to do automatic floor detection in this paper. We aim to in the future combine the approaches of using RSSI and pressure sensors to be able to use crowd sourced data both with and without pressure sensors. In contrast to previous methods, cf. [18], we do not know what floor the WiFis are located or have fingerprints gathered in a calibration step, when using pressure sensors to do automatic floor detection. We do however for now rely on the user to specify the floor at the beginning of a session.

2 Data

Data is collected using Nexus 6 phones, see Fig. 1, using the Combain CPS App available on Google Play, [19]. The data collected consists of individual scans. Each scan has one or more Received Signal Strength Index (RSSI) measurements to the access points in the vicinity. Access points are most commonly WiFis but can also be Bluetooth beacons. A scan also contains an associated time, a pressure value from the barometer and possibly also a GPS location along with an accuracy for the GPS. In the app the user can also specify on which floor he is and if he is inside or not, to be used as ground truth. There is also a possibility to add ground truth positions via a map for measuring the performance of algorithms.

Before exporting the data to the algorithm, access points scanned less than 5 times and scans with less than 3 WiFis are filtered out, to alleviate scan and

[1] Data is available at `https://github.com/SimonEBurgess/WiFi-SLAM`

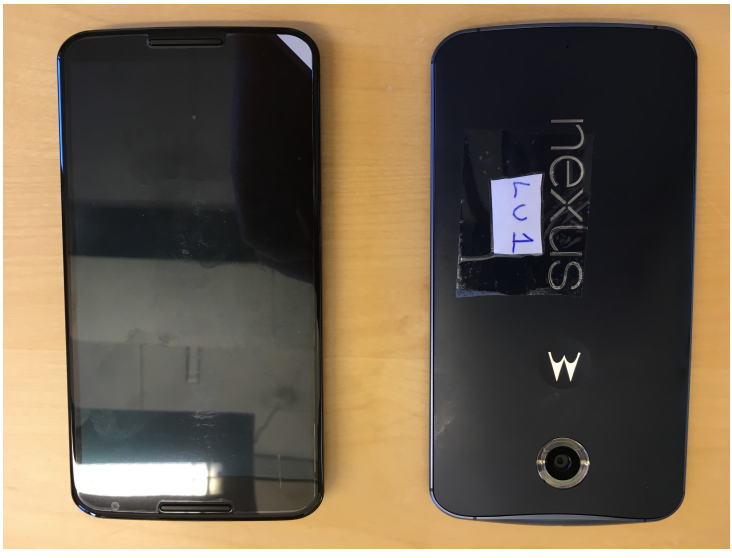


Figure 1: Two Nexus 6 mobile phones used to collect the data. The data used to do simultaneous localization of the smartphone and mapping of the radio environment consists of scans of the RSSI values and pressure readings

access point positions being identifiable. Any access points that can be identified as being non-stationary are removed. This could be e.g. smart devices set to work as access points.

Combain also has an algorithm using a crowdsourced database with positions for access points and scan positions, with a median accuracy of 20m-40m, used for initialization of positions.

3 Methods

Our goal is to build an objective function F to minimize, where the arguments to the objective function include scan positions, access point positions and access point parameters, thus simultaneously locating scan positions, access point positions and parameters. This section is comprised out of three parts; the model used for the signal propagation, the description of the target function to optimize over, and how floor detection is made.

3.1 The model

We model the obtained RSSI measurements, p_{ij}^*, between access point i and scan j, as following the path loss model, with the additional term of adding a linear dampening depending on the number of floors the signal passed through,

$$P_{ij} = C_i - 10\gamma_i \log_{10}(d_{ij}) + n_{ij} C_{floor} + X_{ij}. \quad (1)$$

Here P_{ij} is the model for the obtained RSSI value in dBm, C_i is the measured power at one meter from an access point, d_{ij} is the distance in meters between access point i and scan j, γ_i is the path loss exponent, which is 2 for free space power attenuation, > 2 for power dampening over distances, and < 2 if the signal is enhanced by the environment, n_{ij} is the number of floors separating access point i and scan j, C_{floor} is the floor dampening constant for passing through one floor, and X_{ij} is assumed to be i.i.d. Gaussian noise with zero mean. Let $\mathbf{s}_j = [s_{x,j} s_{y,j} s_{z,j}]^T$ be the 3D position of scan j, i.e. the position at a certain time of a moving receiver device where a scan of the access points is made. Let $\mathbf{t}_i = [t_{x,i} t_{y,i} t_{z,i}]^T$ be the 3D position of access point i. The distance d_{ij} is then

$$d_{ij} = \sqrt{(\mathbf{s}_j - \mathbf{t}_i)^T (\mathbf{s}_j - \mathbf{t}_i)}. \quad (2)$$

3.2 An optimization scheme

Below the different terms in the target function are described, to finally present the full optimization problem.

From the power model we form the sum of squared residuals,

$$F_{pow} = \sum_{i,j} \left(P_{ij} - p_{ij}^* \right)^2. \quad (3)$$

When small, the modeled and measured RSSI values will be close.

We use a set of GPS measurements in the x-y plane, $\mathbf{g}_j$, for a subset J_{GPS} of the receiver indices j. GPS measurements are most commonly obtained outside and at the entrance of the buildings, and close to windows. The accuracy in meters for the horizontal plane, λ_j, is used. A GPS lock in the vertical direction is commonly also obtained, but this has been found too unreliable to use for vertical positioning. The term in the target function pertaining to the GPS measurements

is

$$F_{GPS} = \sum_{j \in J_{GPS}} \frac{1}{\lambda_j} \left(\mathbf{g}_j - \left[s_{x,j} s_{y,j} \right] \right)^T \left(\mathbf{g}_j - \left[s_{x,j} s_{y,j} \right] \right). \quad (4)$$

This term is what locks in the map. Without any absolute GPS-location locks, scan and access point positions could be translated, rotated and mirrored without affecting the term using the RSSI measurements in (3).

From the data, the time stamps of each scan is also available, as well as the ID of the receiver. It is reasonable to assume that receiver positions pertaining to the same device, close in time, should be close spatially. To that purpose, the next term in the target function approximately sums the squared acceleration of every device that has made scans close in time. Let $\mathbf{j} = \{j_1\, j_2\, j_3\}$ be a set of three scan indices that comes from the same receiver device, and has the three indices consecutively in time, and less than ten seconds apart. Let J_{acc} be the set of all such triplets $\mathbf{j}$. The next term F_{acc} is

$$F_{acc} = \sum_{\mathbf{j} \in J_{acc}} (\mathbf{s}_{j_1} - 2\mathbf{s}_{j_2} + \mathbf{s}_{j_3})^T (\mathbf{s}_{j_1} - 2\mathbf{s}_{j_2} + \mathbf{s}_{j_3}). \quad (5)$$

Due to the crude model of the signal environment, and that the GPS locks in the vertical directions are unreliable, the full 3D problem of optimizing over all position coordinates $[s_{x,j} s_{y,j} s_{z,j}]^T$ and $[t_{x,j} t_{y,j} t_{z,j}]^T$ continuously will not yield accurate results. Thus, we limit the vertical position components $s_{z,j}$ and $t_{z,j}$ to be on one of several discrete floors. When a user changes floors in a building, the user commonly do this at certain static locations. It is unlikely that there are several staircases or elevators close to each other. Thus, another term is added, F_Δ, that should be small if floor changes are made close to each other. Let j_1, j_2 be a indices of a pair of consecutive scans from the same user that have different z-coordinates, and J_Δ be the set containing all the pairs of such scan indices. The term F_Δ is then constructed as

$$F_\Delta = \sum_{j_1, j_2 \in J_\Delta} f((\mathbf{s}_{xy,j_1} - \mathbf{s}_{xy,j_2})^T (\mathbf{s}_{xy,j_1} - \mathbf{s}_{xy,j_2})), \quad (6)$$

where $\mathbf{s}_{xy,j} = \left[s_{x,j} s_{y,j} \right]$, $f(\cdot)$ is a identity function if $\left[s_{x,j_1} s_{y,j_1} \right]$ are within 15m of $\left[s_{x,j_2} s_{y,j_2} \right]$, and constant 0 otherwise. Thus, a small F_{floor} will have scan positions

from users changing floors close to each other in the x-y-plane, but not affected if the floor change is far away from other floor changes.

Our full target function F using (3), (4), (5), and (6) is constructed as

$$F = F_{pow} + \lambda_1 F_{GPS} + \lambda_2 F_{acc} + \lambda_3 F_{\Delta}, \tag{7}$$

where $\lambda_1,\ \lambda_2,\ \lambda_3$ are scaling parameters > 0 that determines how important each part is in the target function.

The target function F is to be optimized over the scan positions $[s_{x,j}s_{y,j}s_{z,j}]$, access point positions $[t_{x,i}t_{y,i}t_{z,i}]$, access point parameters C_i, γ_i, and floor dampening C_{floor}. Of these, the z-axis coordinates $s_{z,j}, t_{z,i}$ are discrete heights for each floor, and the rest are in $\mathbb{R}$. Thus, we have a mix of continuous and discrete variables. Holding the floor positions constant, we use a Levenberg-Marquardt like algorithm, [20], to find a local optimum of (7).

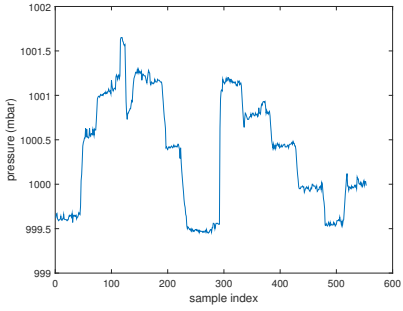
It remains to find good initializations. For x- and y- positons, the Combain database is used to get rough coordinates in the vicinity of the building. The parameters for the access points are all initialized as $C_i = -37dB$ and $\gamma_i = 2.5$. Floor dampening is initialized as $C_{floor} = -15dBm$.

To optimize over all permutations of discrete floor positions for scans and access points is not feasible. Thus, an estimate of floor levels needs to be done. In [17] a clustering approach looking at similarities between the RSSI measurements in each scan is used, but is found to be insufficient as the only means to find floor level.

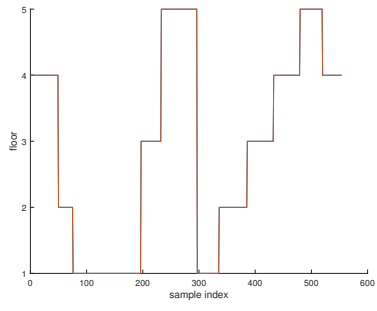
3.3 Floor estimation

We here describe how the smartphone pressure sensor is used to estimate the floor each scan position is on. Each device used in a data set generally produce scans close to each other in time, see Section 2. Here we look at the pressure from these consecutive scans to determine floor changes.

Air pressure is commonly used in conjunction with meteorological data to determine height in e.g. altimeters in airplanes. As access to real time meteorological data is not feasible in the current use case, air pressure can not be used to explicitly determine exact height, and the precision of such altimeter data is generally not exact enough for the problem of determining floor level. It is possible, however, to study the change in air pressure to try and determine when a floor change takes place.



(a) Pressure data from consecutive scans



(b) Ground truth floors for consecutive scans

Figure 2: The measured pressure compared to the ground truth floor for consecutive scans for one device.

In Fig. 2, the air pressure over time is compared to when the user has ascended or descended one or several floors. It is clear that fast pressure changes are often the result of a floor change. When the user ascends a stair the pressure drops and vice versa. The signal is quite noisy, and may contain pressure changes like the one around sample index 120 in Fig. 2, which could be the result of winds, the user opening and closing a door in a confined space or some other source. Thus we need a method to determine which changes in pressure actually should signal a floor change.

The pressure signal is convolved with the derivative of a Gaussian kernel to detect strong changes in pressure. See Fig. 3. Each such strong change is then considered to be a possible floor change. To compensate for drift due to weather changes, a linear trend removal is done between possible floor changes where sufficient time has passed. See Fig. 4. As the floor spacing is unknown, the signal is not discretized yet, instead an early estimate of the change in altitude is calculated by comparing the median pressures of neighbouring intervals.

For a naive approach of which height the user has been on at each sample, consecutive height changes are added. See Fig. 5. When these results are compared to the ground truth in Fig. 2 b), it is clear that some drift has occurred. When the user first arrives at floor 5, this is estimated to be very close to the height that was on the 4th floor at the beginning of the session.

The spacing between each floor, h, is assumed to be uniform in each separate

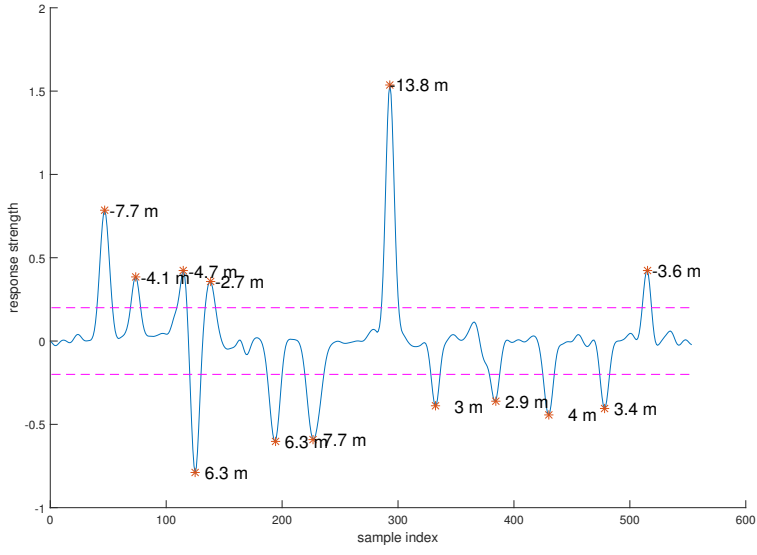


Figure 3: Result of pressure signal convolved with the derivative of a Gaussian kernel. A threshold was added after studying several scenarios. Any pressure change stronger than that threshold is hypothesized to be a floor change. Note that the delta height was not obtained from the response strength but is included for clarity.

building. It is determined from the difference in heights, d_i by solving

$$h^* = \underset{h \in [2.5\ 5]}{\operatorname{argmin}} \sum |d_i - round(d_i/h) \cdot h| \,. \tag{8}$$

Once the floor spacing has been estimated, every difference in height is discretized into a number of floor changes. Note that we assume here that the user in the beginning of the session enters the starting floor number. We expect to be able to do this initialization automatically in the future; combining the GPS accuracy and positions one should be able to determine when in the data set a user enters a building.

In Fig. 6 the resulting floor estimations are visualized along with ground truth. There is a faulty detection at sample indices 120-130. The user went

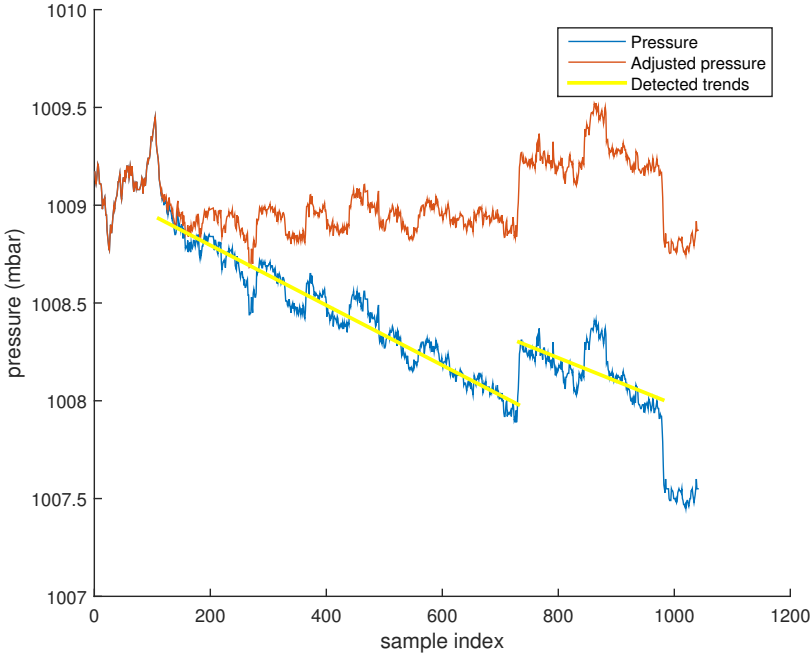


Figure 4: A pressure signal with a clear trend due to weather changes (blue). The trend is removed with a linear regression for separate sub intervals, detected using the hypothesized floor changes.

outside for some time which led to noise in the signal. The noise might have come from the doors opening and closing, wind outside, uneven ground or other sources. Whatever the cause in this case, we have found that outdoor pressure readings generally contain a lot more noise than indoor readings. It is therefore desirable to detect when a user travels outside and simply disregard all pressure changes that take place while outdoors. To determine when the user is outdoors we studied the current accuracy of the GPS signal to the phone that was used to perform the readings. The GPS accuracy has a strong connection to the signal strength from the GPS satellites, if the receiver has LOS or a multipath component to the GPS satellites, the number of satellites observed, and the dilution of precision. The GPS accuracy should thus be improved when outdoors.

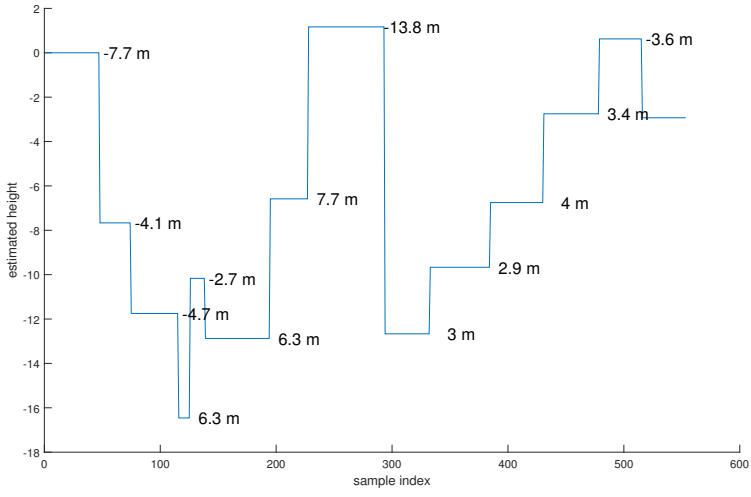


Figure 5: A first estimate of heights of the scan positions using pressure sensor data

The GPS accuracy also contains noise and we therefore smooth the signal with a Gaussian kernel before comparing it to a threshold that was empirically selected. When the filtered signal is below the threshold we determine that the user has been outdoors. This method will signal some false positives if the user is indoors, but in a place where GPS signals can permeate efficiently, e.g. in the vicinity of large windows. One way to deal with this issue is to check the interval length and GPS positions of the hypothesized outdoor part and dismissing it if it is not long enough in time or distance. Another way is to study the interval between two height changes and perform a majority vote among scan indices to decide whether the interval is indoors or outdoors. We use a combination of these methods to get a final result, see Fig. 7.

When the floors of the scan positions has been initiated, the floors of access points are then set to the same as the floor of the scan with the strongest RSSI value to it.

The floor detection method above suffers from the issue of an erroneous detection in the beginning sometimes will propagate to the scans that come after. E.g. an erroneous detection of a floor increase in the beginning sometimes makes

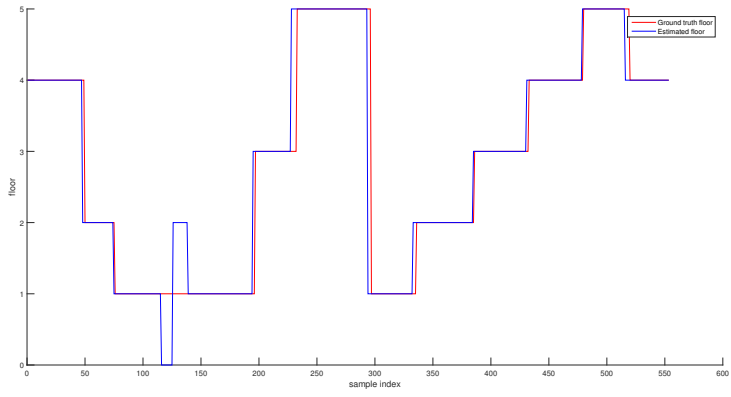


Figure 6: The calculated floors (blue) compared to the ground truth (red) for each scan position for consecutive scans

all the scans afterwards to be detected as one floor higher than they are. To alleviate this, a final step is proposed where the optimization in Section 3.2 and floor detection are iterated. This floor estimation is described below.

1. Use pressure sensor where available to get a floor estimate using the methods above in this section.

2. Optimize over all positions and access point paramaters using the estimated floors and the methods in Section 3.2.

3. For each floor, check for 5x5m^2 tiles where scans are similar to each other inside the tiles. These tiles will become representative of that floor.

4. For each interval, each scan in that interval will compare how similar it is to each tile. If similar enough, it will cast a vote for that tile, and by that cast a vote to which floor it belongs to.

5. Majority vote determines which floor each interval belongs to. This gives what floor each scan belongs to.

6. Floors of access points are set to the same as the floor of the scan with the strongest RSSI value to it.

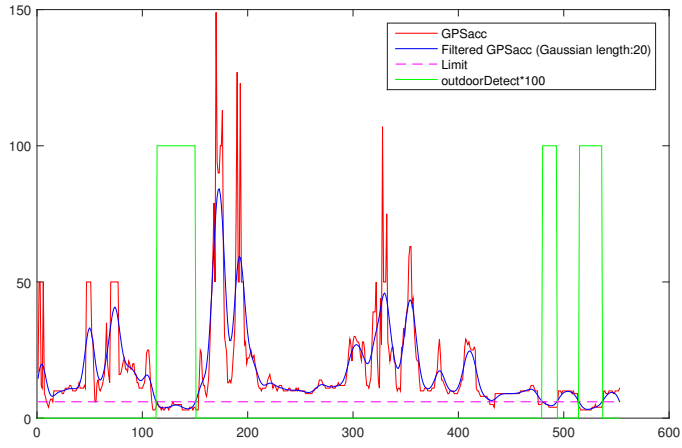


Figure 7: For each scan index, the high level green curve shows the scan indices detected as outdoors, using the GPS accuracy (red) and the smoothed GPS accuracy (blue). The ground truth outdoors are located between scan index 110 and 150.

Steps 2-6 are iterated until convergence, which is for our two data sets once and twice. The final floor estimates of scans and access points is then used together with the optimization in Section 3.2 to produce the final results. Although we do not have such data here, scans that do not belong to any pressure interval, or do not have pressure data, can also be set to a floor by this method, as long as the data has some partial sequential pressure readings.

4 Results

The method has been tested in two buildings on the Lund University Campus. The buildings are called Centre for Mathematical Sciences (M.C) and E-house.

M.C. is an approximately 130x35m building mixed with mainly lecture halls in the lower floors and office space in the upper floors with the office spaces on the sides of a long corridor. The data was collected in five separate sessions and contains 3070 scans in total. 606 different access points were seen by the scans, with 107 of them were manually found inside. However, more are expected to be

inside. The 107 found are used as ground truth positions for access points.

E-house is an approximately 150x100m building mixed with office, class rooms and laboratories, consisting of five floors plus one cellar floor. The data was collected in four separate sessions and contains 2770 scans in total. 793 different access points were seen by the scans, with 21 of them were manually found inside. However, more are expected to be inside. The 21 found are used as ground truth positions for access points.

A Nexus 6 smartphone was used as a receiver, running the Combain CPS app, generating a scan usually once every 3-10 seconds. Ground truth was added manually by the users in the app while performing the scans, at about 0.5-1% of the scan points by selecting the current position on a satellite map. It should be noted that this method does not produce a perfect ground truth due to both the difficulty of selecting a perfect position on a crude map as well as the possible errors of the map itself. We estimate that the user precision is a couple of meters. The users entered ground truth floor at every floor change to allow us to compare it to the floor estimations. Different users performed the different sessions.

Table 1: Floor Estimation Results

Building	Correct floor	Ratio
M.C.	2913/3070	94.9%
E-house	2706/2770	97.7%

It should be noted that when a user enters ground truth of floor changes in the app, some users may enter the change while entering the staircase, some users in the middle and some users when exiting. Some users may even forget to make the change until they have taken a few steps away from the stairs. This should be taken into consideration when judging the scores of Table 1. A more fair view of the results can be achieved by studying a plot comparing ground truth and estimates of floors as depicted in Fig. 8 and 9b.

By studying Fig. 8 and 9a it can be determined that the floor estimations work very well for M.C. but performs a lot worse in E-house. The poor results of E-house are due to a false positive floor change detection early in the first session. That error carries over to all subsequent scans of that session, until there is a new session where the user once again enters ground truth floor as initialization value. By running Step 2-6 in Section 3.3 two times, the floor detection converged to the results in Table 1, visualized in Fig. 9b.

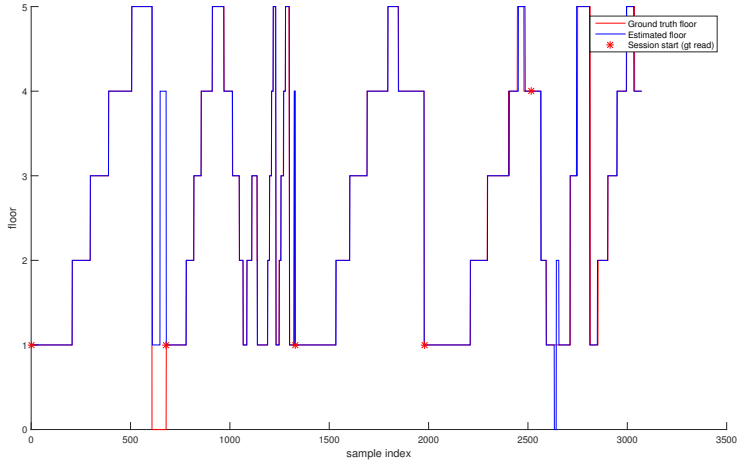


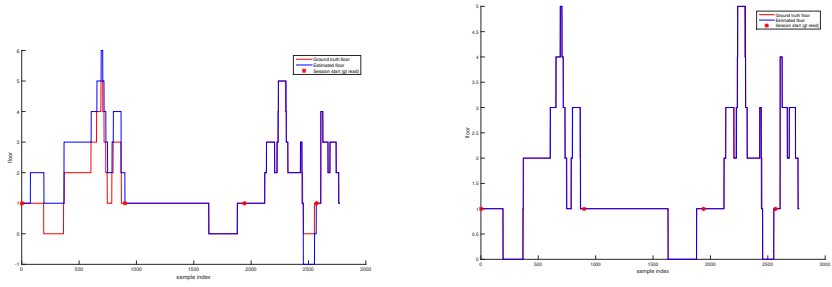
Figure 8: Results for floor detection for all sessions in M.C. Ground truth floors for the scans is in red. Estimated floor is in blue. Beginnings of new sessions are marked by a red star in the plot.

Table 2: The horizontal distance error of smartphone and access point positions

	WiFi error		Scan errors	
Building	Mean	Median	Mean	Median
M.C.	11.2 m	7.16 m	13.0m	12.5 m
E-house	21.9 m	19.4 m	21.5 m	16.4 m

For the final results the known positions of each access point has been compared to its estimate and the available ground truth scan positions have been compared to the estimated scan positions. In order to remove any positive bias, ground truth scan positions with GPS accuracy < 10m is removed. This mainly coincides with removing any comparison to ground truth positions gathered by the user while moving outside. The parameters in (7) are set to $\lambda_1 = 10, \lambda_2 = 2, \lambda_3 = 10$. These are chosen heuristiacly, manually trained on a different data set for 2D used in our technical report, [17].

In Table 2, the mean and median error in the horizontal direction for scan and access point positions can be seen.



(a) Initial floor detection results

(b) Final floor detection results

Figure 9: Floor detection results in M.C. Ground truth floors for the scans is in red. Estimated floor is in blue. Beginnings of new sessions are marked by a red star in the plot. a) shows initial results for floor detection, before running Step 2-6 in Section 3.3 b) shows final results for floor detection.

In Fig. 10 and 11 the scans from all sessions are displayed in 3D plots. Note that it is almost possible to visually determine the quality of the estimations by comparing the tracks in the building. In M.C. you can easily determine where hallways are on the top three floors while in E-house, which did not have as high quality results, does not have nearly as neatly defined paths. We hypothesize that, when access to more crowdsourced data, the quality will improve due to the amount of data that will be produced. This assumption is based on our observations of the effects of increasing the amount of data to perform the calculations on. This can be seen if comparing Fig. 10 and 12 where the former has data from five sessions and the latter only from a subset of that data. In Fig. 12 the tracks are even outdoors on floors 2 to 5 more often than not while in Fig. 10 they are much more confined inside the walls of the building. We hypothesize that more data in E-House will also increase performance, as the bulding has a more elaborate floor plan than M.C. Also note that in no way has the building outline been used in the calculations, although this is a possibility for the future. For this smaller data set in M.C. we have horizontal errors as follows: Mean WiFi error of 17.6m, median WiFi error 15.5m, mean scan error of 15.4m, and median scan error of 13.4m. Compared to the results for M.C. in Table 2, where the full data set is used, it is notably worse.

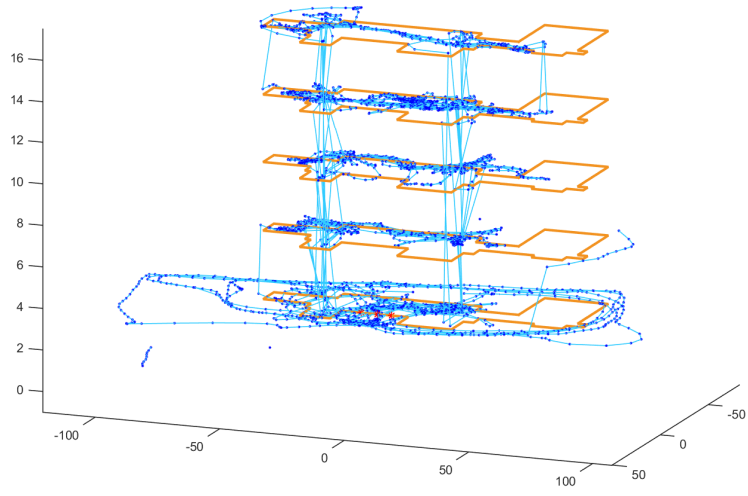


Figure 10: 3D plot of all calculated smartphone positions in and around M.C. Smartphone positions, i.e. scan positions, are blue dots, with the lines connecting scans that are close in time to each other to show trajectories of users moving. The building outline is only added for visualization purposes, and has not been used in the method. Note how one can spot where the stairs are located in between floors.

5 Conclusions

We have presented a method for simultaneously mapping the radio environment and positioning several smartphones in multi-story buildings. There is no communication between the phones, and the localization is done offline using the data collected from all users, simultaneously optimizing model parameters of the scan positions, WiFi positions, and WiFi parameters. No prior calibration of the radio environment, the floor plan, or added infrastructure has been used. Only signal strength from preexisting WiFi and Bluetooth beacons in the buildings have been used, together with pressure sensors in the smartphones.

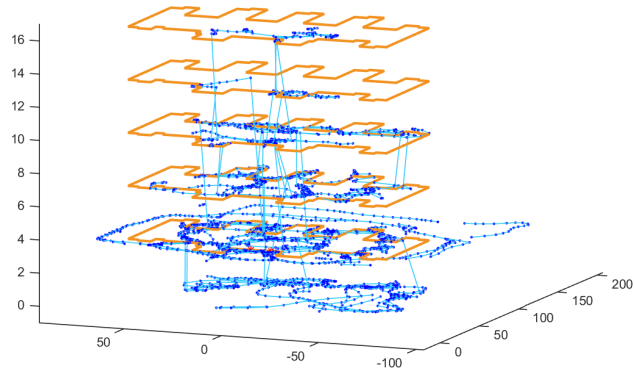


Figure 11: 3D plot of all calculated smartphone positions in and around E-house. Scan positions, i.e. smartphone positions, are blue dots, with the lines connecting scans that are close in time to each other to show trajectories of users moving. The building outline is only added for visualization purposes, and has not been used in the method. Note that the cellar movement has been detected.

Comparing with our previous poster abstract, [17], we determine the floor level by using the pressure sensors instead of manually entering the floor level. The target function for the optimization also include terms for clustering floor changes, i.e. staircase and elevator positions, and dealing with virtual access points having the same physical access points. The mean error in the x-y-plane of the smartphones has increased from 15.2m in our technical report to 21.9m in the E-House, whereas in the M.C. building the mean error x-y-plane has dropped from 13.5 in our technical report to 11.2m. The difference in performance is most likely due to different data sets. The previous data sets only had three floors for both respective buildings, whereas we here present data sets with five floors. New data sets have been used as pressure data is a crucial part of our algorithms to automatically detect floor level.

In a similar approach in [15], a median error of 7m for the smartphone is calculated for a large building. They do however only consider single-story buildings, and use manually added GPS signals inside the building. The floor detection

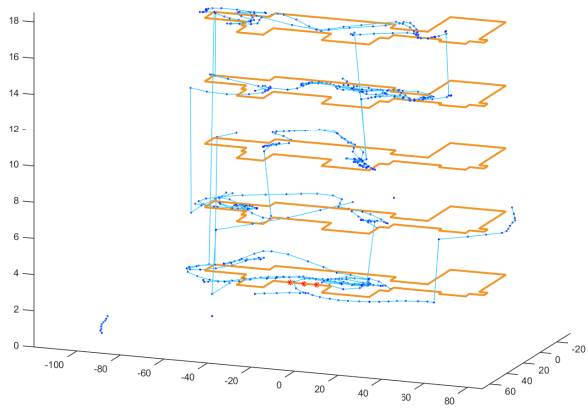


Figure 12: 3D plot of only using a smaller subset of the data in and around M.C. Scan positions, i.e. smartphone positions, are blue dots, with the lines connecting scans that are close in time to each other to show trajectories of users moving. The building outline is only added for visualization purposes, and has not been used in the method. Note the deteriorating paths, compared to Fig. 10 where all data is used.

has an overall accuracy of 96% compared with the ground truth gathered by users, compared to the FCC rules for mobile carriers, [16], which states that an accuracy of 3m vertically in 67% of the cases is needed.

Future work includes merging data gathered from phone both with pressure sensors and without, and reliably automatically detecting ground floor using GPS positioning. Furthermore, a genetic algorithm for the mixed continuous and discrete optimization problem would be of great use. It could help minimize the risk of local optima, alleviate the dependency on initialization and to evaluate how effective the target function is.

References

[1] Charalampos Papamanthou, Franco P. Preparata, and Roberto Tamassia, *Algorithmic Aspects of Wireless Sensor Networks: Fourth International Workshop, ALGOSENSORS 2008, Reykjavik, Iceland, July 2008. Revised Selected Papers*, chapter Algorithms for Location Estimation Based on RSSI Sampling, pp. 72–86, Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.

[2] Binghao Li, James Salter, Andrew G. Dempster, and Chris Rizos, "Indoor positioning techniques based on wireless lan," in *First IEEE International Conference on Wireless Broadband and Ultra Wideband Communications*, 2006, pp. 13–16.

[3] M. Bshara, U. Orguner, F. Gustafsson, and L. Van Biesen, "Fingerprinting localization in wireless networks based on received-signal-strength measurements: A case study on wimax networks," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 283–294, Jan 2010.

[4] P. Vorst, J. Sommer, C. Hoene, P. Schneider, C. Weiss, T. Schairer, W. Rosenstiel, A. Zell, and G. Carle, "Indoor positioning via three different rf technologies," in *RFID Systems and Technologies (RFID SysTech), 2008 4th European Workshop on*, June 2008, pp. 1–10.

[5] Anton Schwaighofer, Marian Grigoras, Volker Tresp, and Clemens Hoffmann, "Gpps: A gaussian process positioning system for cellular networks," in *Advances in Neural Information Processing Systems 16*, S. Thrun, L. K. Saul, and B. Schölkopf, Eds. MIT Press, 2004.

[6] J. Shang, X. Hu, F. Gu, D. Wang, and S Yu, "Improvement schemes for indoor mobile location estimation: A survey," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[7] J. Huang, D. Millman, M. Quigley, D. Stavens, S. Thrun, and A. Aggarwal, "Efficient, generalized indoor wifi graphslam," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, May 2011, pp. 1038–1043.

[8] Zheng Yang, Chenshu Wu, and Yunhao Liu, "Locating in fingerprint space: Wireless indoor localization with little human intervention," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, 2012, Mobicom '12, pp. 269–280.

[9] He Wang, Souvik Sen, Ahmed Elgohary, Moustafa Farid, Moustafa Youssef, and Romit Roy Choudhury, "No need to war-drive: Unsupervised indoor localization," in *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services*, 2012, MobiSys '12.

[10] Anshul Rai, Krishna Kant Chintalapudi, Venkata N. Padmanabhan, and Rijurekha Sen, "Zee: Zero-effort crowdsourcing for indoor localization," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, New York, NY, USA, 2012, Mobicom '12, pp. 293–304, ACM.

[11] Chenshu Wu, Zheng Yang, Yunhao Liu, and Wei Xi, "Will: Wireless indoor localization without site survey," *IEEE Transactions on Parallel and Distributed Systems*, vol. 24, no. 4, pp. 839–848, 2013.

[12] Anthony LaMarca, Jeff Hightower, Ian Smith, and Sunny Consolvo, *UbiComp 2005: Ubiquitous Computing: 7th International Conference, UbiComp 2005, Tokyo, Japan, September 11-14, 2005. Proceedings*, chapter Self-Mapping in 802.11 Location Systems, Springer Berlin Heidelberg, 2005.

[13] Daniel Turner, Stefan Savage, and Alex C. Snoeren, "On the empirical performance of self-calibrating wifi location systems," in *Proceedings of the 2011 IEEE 36th Conference on Local Computer Networks*, 2011, LCN '11.

[14] Abhishek Goswami, Luis E. Ortiz, and Samir R. Das, "Wigem: A learning-based approach for indoor localization," in *Proceedings of the Seventh COnference on Emerging Networking EXperiments and Technologies*, 2011, CoNEXT '11, pp. 3:1–3:12.

[15] Krishna Chintalapudi, Anand Padmanabha Iyer, and Venkata N. Padmanabhan, "Indoor localization without the pain," in *Proceedings of the Sixteenth Annual International Conference on Mobile Computing and Networking*, 2010.

[16] Federal Communications Commission, "Wireless e911 location accuracy requirements," 2015.

[17] Simon Burgess, Björn Lindquist, Rasmus Ljungberg, Koustubh Sharma, and Kalle Åström, "Indoor localization using smartphones in multi floor environments without prior calibration or added infrastructure," in *Proceedings of the Sixth International Conference on Indoor Positioning and Indoor Navigation*, 2015, IPIN 2015.

[18] P. Gupta, S. Bharadwaj, S. Ramakrishnan, and J. Balakrishnan, "Robust floor determination for indoor positioning," in *Communications (NCC), 2014 Twentieth National Conference on*, 2014.

[19] Combain AB, "Combain CPS," `https://play.google.com/store/apps/details?id=com.combain.cpsdemo`, 2016.

[20] Donald W. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the Society for Industrial and Applied Mathematics*, 1963.