A CONVEX RELAXATION FOR APPROXIMATE MAXIMUM-LIKELIHOOD 2D SOURCE LOCALIZATION FROM RANGE MEASUREMENTS

Pınar Oğuz-Ekim, João Gomes, João Xavier and Paulo Oliveira

Institute for Systems and Robotics-Instituto Superior Tecnico
Department of Electrical and Computer Engineering, Lisboa, Portugal
{*poguz, jpg, jxavier, pjcro*}*@isr.ist.utl.pt*

ABSTRACT

This paper addresses the problem of locating a single source from noisy range measurements in wireless sensor networks. An approximate solution to the maximum likelihood location estimation problem is proposed, by redefining the problem in the complex plane and relaxing the minimization problem into semidefinite programming form. Existing methods solve the source localization problem either by minimizing the maximum likelihood function iteratively or exploiting other semidefinite programming relaxations. In addition, using squared range measurements, exact and approximate least squares solutions can be calculated. Our relaxation for source localization in the complex plane (SLCP) is motivated by the near-convexity of the objective function and constraints in the complex formulation of the original (non-relaxed) problem. Simulation results indicate that the SLCP algorithm outperforms existing methods in terms of accuracy, particularly in the presence of outliers and when the number of anchors is larger than three.

***Index Terms*—** Single source localization, maximum likelihood estimation, nonconvex and nonsmooth minimization, semidefinite programming.

1. INTRODUCTION

The problem of estimating the position of a source using only the distances to a set of points with known coordinates (called anchors) appears in many engineering and scientific applications. This problem has been solved by different schemes and using various types of range-like measurements such as time of arrival (TOA) or received signal strength (RSS) [1, 2, 3, 4, 5, 6].

Maximum likelihood (ML) based solutions are of particular interest due to their asymptotically optimal performance. However, the ML estimator requires the minimization of a nonconvex and nonsmooth cost function. One approach for finding a local solution of the ML problem is through the use of iterative minimization techniques [1]. In [1], the authors exploited the special structure of the cost function to derive a fixed-point iterative scheme and another method where each

iteration consists of solving a generalized trust region sub-problem (GTRS) by using TOA based range measurements. These algorithms might fail to find a global solution and have either slow convergence rate or high computational cost. In [2], the ML problem based on TOA range measurements was reformulated and relaxed to construct a suboptimal but simpler optimization problem. The resultant relaxed ML problem is a semidefinite program (SDP), and therefore convex. However, as discussed in [3], the optimal solution of this relaxed SDP does not always satisfy the near-rank-1 constraints of acceptable solutions to the source localization problem.

An alternative approach is to define the source localization problem as a least squares (LS) problem using squared ranges (SR) or SR difference TOA based measurements. Despite its nonconvexity, the SR-LS problem can be solved globally and efficiently by resorting to GTRS [3]. The SR-LS approaches are computationally simpler than iterative minimization algorithms but they provide less accurate solutions than those provided by ML approaches [3], because they are suboptimal in the ML sense. Besides, SR-LS solutions are known to undergo significant degradation in the presence of outliers [7], which commonly affect practical range measurement systems.

The SLCP and techniques described up to this point are centralized methods. However, it can be interesting to solve the source localization problem in a distributed manner in wireless sensor networks. In [5], the authors proposed a distributed incremental gradient method to solve the ML problem using RSS measurements. Nevertheless, like most iterative methods applied to nonconvex problems, it may get trapped in local minima. The problem can also be formulated as a convex feasibility problem and solved via distributed projection onto convex sets (POCS) using RSS measurements [6]. POCS converges to a limit point or to a limit cycle in the vicinity of the true source position.

In this paper, we consider centralized ML source localization because of its asymptotically optimal characteristics and superiority compared to SR-LS in the presence of outliers [7]. Note that our ML function is built under the assumption of Gaussian noise, which leads to a LS estimator problem. We reformulate the ML problem as a two stage optimization problem, given that the problem has two sets of variables. In the first step, one of the variable sets is fixed and the unconstrained optimization part of the problem is solved

with respect to the other variable set. In the second step, constraints are described in the complex plane and the optimization problem is relaxed to an SDP. This relaxation was found to be more accurate than others which have been previously published [2]. That behavior is heuristically interpreted in Section 2 based on the convexity of the cost function and the near-convexity of the constraint set. The accuracy of SLCP makes it a convenient initialization method for iterative refinement methods, which therefore require fewer iterations to converge and/or are less likely to be trapped in undesirable local extrema of the ML cost function.

The organization of this paper is as follows. In Section 2, ML location estimation is formulated and the SLCP algorithm is introduced. In Section 3, simulation results for SLCP are presented and its performance is benchmarked against other methods. Finally, conclusions are drawn in Section 4. Vectors and matrices are denoted by boldface lowercase and uppercase letters, respectively. The i-th component of a vector $\mathbf{a}$ is written as a_i. The superscript T (H) denotes the transpose (Hermitian) of the given real (complex) vector or matrix. Below, I_m is the $m \times m$ identity matrix and the $\mathbf{1}_m$ is the vector of m ones. For symmetric matrix $\mathbf{X}$, $\mathbf{X} \geq 0$ means that $\mathbf{X}$ is positive semidefinite.

2. PROBLEM FORMULATION

Let $x \in \mathbb{R}^2$ be the unknown source position, $a_i \in \mathbb{R}^2$ be known sensor positions (anchors) for $i = 1, .., m$, and $r_i = \|x - a_i\| + w_i$ be the measured range between the i-th anchor and the source, contaminated by zero mean Gaussian noise w_i with variance σ^2. The ML source localization problem is equivalent to

$$\underset{x}{\text{minimize}} \quad f(x) = \sum_{i=1}^{m}(\|x - a_i\| - r_i)^2. \tag{1}$$

The objective function in (1) is nonconvex and nonsmooth. If we examine the terms $\|x - a_i\|$ and r_i, the difference between the true range and observed range is actually equivalent to the distance between the source position and the circle with center a_i and radius r_i, i.e., $\|\|x - a_i\| - r_i\| = \|x - y_i\|$, where y_i is located in the intersection of the line connecting a_i and x with the circle $\{z : \|z - a_i\| = r_i\}$ (Fig. 1).
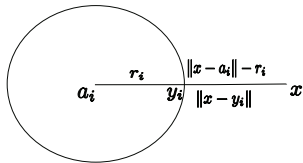

Fig. 1. Geometric interpretation of terms in ML cost function.

An equivalent formulation is therefore,

$$\underset{x, y_i}{\text{minimize}} \quad \sum_{i=1}^{m}\|x - y_i\|^2 \tag{2}$$
$$\text{subject to} \quad \|y_i - a_i\| = r_i.$$

If we fix y_i, the solution of (2) with respect to x is an unconstrained optimization problem whose solution is obtained by invoking the optimality conditions

$$\frac{\partial f(x)}{\partial x} = \sum_{i=1}^{m}(x - y_i) = 0 \Rightarrow x^* = \bar{y} = \frac{y_1 + .. + y_m}{m}.$$

Now, (2) becomes a constrained variance minimization problem

$$\underset{y_i}{\text{minimize}} \quad \sum_{i=1}^{m}\|\bar{y} - y_i\|^2 \tag{3}$$
$$\text{subject to} \quad \|y_i - a_i\| = r_i.$$

Geometrically, the constraints of (3) define circle equations, which can be compactly described in the complex plane as

$$\underset{y_i, \phi_i}{\text{minimize}} \quad \sum_{i=1}^{m}\|\bar{y} - y_i\|^2 \tag{4}$$
$$\text{subject to} \quad y_i = a_i + r_i e^{j\phi_i}.$$

Invoking the centering operator $\Pi = (\mathrm{I}_m - \frac{1}{m}\mathbf{1}_m\mathbf{1}_m^T)$, which subtracts the mean of a vector from each of its components, (4) is rewritten in matrix form as

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \quad (\mathbf{a} + \mathbf{R}\boldsymbol{\theta})^H \Pi(\mathbf{a} + \mathbf{R}\boldsymbol{\theta}) \tag{5}$$
$$\text{subject to} \quad |\theta_i| = 1,$$

where $\mathbf{a} = \begin{bmatrix} a_1 \cdots a_m \end{bmatrix}^T \in \mathbb{C}^m, \mathbf{R} = \mathrm{diag}(\begin{bmatrix} r_1 \cdots r_m \end{bmatrix}) \in \mathbb{R}^{m \times m}$ and $\boldsymbol{\theta} = \begin{bmatrix} e^{j\phi_1} \cdots e^{j\phi_m} \end{bmatrix}^T \in \mathbb{C}^m$. Expanding the objective function in (5) and deleting the constant terms yields

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \quad 2\,\mathrm{Re}(\mathbf{a}^H\Pi\mathbf{R}\boldsymbol{\theta}) - \frac{1}{m}\boldsymbol{\theta}^H\mathbf{R}\mathbf{1}_m\mathbf{1}_m^T\mathbf{R}\boldsymbol{\theta} \tag{6}$$
$$\text{subject to} \quad |\theta_i| = 1.$$

Define $\mathbf{R}^H\mathbf{1} = \mathbf{r}$ and $\mathbf{a}^H\Pi\mathbf{R} = \mathbf{c}^H$. Thus,

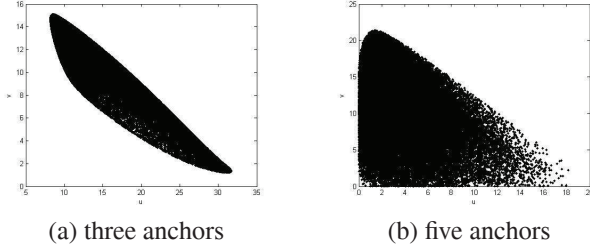$$\underset{\boldsymbol{\theta}}{\text{minimize}} \quad 2\,\mathrm{Re}(\mathbf{c}^H\boldsymbol{\theta}) - \frac{1}{m}\boldsymbol{\theta}^H\mathbf{r}\mathbf{r}^T\boldsymbol{\theta} \tag{7}$$
$$\text{subject to} \quad |\theta_i| = 1.$$

The value of $\mathrm{Re}(\mathbf{c}^H\boldsymbol{\theta})$ is in the interval $[-|\mathbf{c}^H\boldsymbol{\theta}|, |\mathbf{c}^H\boldsymbol{\theta}|]$, and for any $\boldsymbol{\theta}$ we can always find an auxiliary angle γ such that the rotated vector $\boldsymbol{\theta}_0 = \boldsymbol{\theta}e^{j\gamma}$ satisfies $\mathrm{Re}(\mathbf{c}^H\boldsymbol{\theta}_0) = -|\mathbf{c}^H\boldsymbol{\theta}_0|$ without affecting the other term in the cost function or the constraints. With this modification, the optimization problem (7) becomes easier to handle.

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \quad -2|\mathbf{c}^H\boldsymbol{\theta}| - \frac{1}{m}\boldsymbol{\theta}^H\mathbf{r}\mathbf{r}^T\boldsymbol{\theta} \tag{8}$$
$$\text{subject to} \quad |\theta_i| = 1.$$

Before relaxing (8) we further investigate the geometric properties of this formulation by introducing new variables u and v and writing the optimization problem as

$$\underset{u, v}{\text{maximize}} \quad 2\sqrt{u} + \frac{1}{m}v$$
$$\text{subject to} \quad \{\exists \boldsymbol{\theta} : |\theta_i| = 1; u = \boldsymbol{\theta}^H\mathbf{c}\mathbf{c}^H\boldsymbol{\theta}, v = \boldsymbol{\theta}^H\mathbf{r}\mathbf{r}^T\boldsymbol{\theta}\}. \tag{9}$$

(a) three anchors   (b) five anchors

Fig. 2. The constraint set (u, v) for randomly generated $\boldsymbol{\theta}$ that satisfies $|\theta_i| = 1$ for three and five anchors.

The objective function in (9) is concave with respect to u and v. Fig. 2 depicts points in the constraint set (u, v) for randomly generated $\boldsymbol{\theta}$ that satisfies $|\theta_i| = 1$ and $m = 3$ or 5 anchors. As seen from the figure, these constraint sets are *almost convex*, such that (9) describes an "*almost convex*" optimization problem. To motivate our proposed relaxation we hypothesize that dropping a subset of constraints which destroy convexity should not severely disrupt the solution of the problem. To this end (8) is expanded and rewritten as

$$\begin{aligned} \underset{\boldsymbol{\theta}}{\text{maximize}} \quad & 2\sqrt{\text{tr}(\mathbf{c}\mathbf{c}^H\boldsymbol{\theta}\boldsymbol{\theta}^H)} + \tfrac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\boldsymbol{\theta}\boldsymbol{\theta}^H) \\ \text{subject to} \quad & |\theta_i| = 1. \end{aligned} \quad (10)$$

Following standard manipulations we introduce the new variable $\boldsymbol{\Phi} = \boldsymbol{\theta}\boldsymbol{\theta}^H$ and an associated rank constraint,

$$\begin{aligned} \underset{\boldsymbol{\Phi}}{\text{maximize}} \quad & 2\sqrt{\text{tr}(\mathbf{c}\mathbf{c}^H\boldsymbol{\Phi})} + \tfrac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\boldsymbol{\Phi}) \\ \text{subject to} \quad & \boldsymbol{\Phi} \geq 0 \\ & \Phi_{ii} = 1 \\ & \text{rank}\,\boldsymbol{\Phi} = 1. \end{aligned} \quad (11)$$

Finally, a relaxed SDP formulation is obtained by introducing the epigraph variable t and dropping the rank constraint,

$$\begin{aligned} \underset{\boldsymbol{\Phi}, t}{\text{maximize}} \quad & t + \tfrac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\boldsymbol{\Phi}) \\ \text{subject to} \quad & \boldsymbol{\Phi} \geq 0 \\ & \Phi_{ii} = 1 \\ & 4\mathbf{c}^H\boldsymbol{\Phi}\mathbf{c} \geq t^2. \end{aligned} \quad (12)$$

Remark that the solution of the optimization problem (12) is a positive semidefinite matrix, hopefully with near-1 rank. Afterwards, the source coordinates are estimated by singular value decomposition (SVD) of $\boldsymbol{\Phi}$ [4].

3. SIMULATIONS AND COMPARISONS WITH THE EXISTING METHODS

In this section we will demonstrate the performance of the SLCP algorithm in simulation and compare it with other existing methods from [1, 3]. In [3], the authors consider LS approaches for locating a source from SR measurements, which is termed SR-LS. The problem is cast as minimizing

a quadratic function subject to a single quadratic constraint, which is efficiently solved by GTRS. Despite a fundamental mismatch with the likelihood function, that method is frequently more accurate than the ML relaxation of [2], which may fail to find a valid solution satisfying rank-1 constraints. In this section, however, we will see that as a relaxed ML method, SLCP outperforms SR-LS for more than three anchors. The difference in performance is more significant in the presence of outliers.

We also use SLCP as initialization of iterative methods, in which case its impact on the number of required iterations for convergence is of concern as well. We have used the simple fixed point (SFP) method described in [1]. SFP is derived by using optimality conditions and mimicking the Weiszfeld method which is used to solve the localization problem to minimize the sum of euclidean distances. We used $\|\nabla f(x^k)\| \leq 10^{-7}$ as a stopping criterion for SFP, where $\nabla f(x^k)$ is the gradient of the objective function in (1) at k-th iteration.

Example 1: In this example, SLCP is compared with SR-LS using five anchors. We performed 1000 Monte Carlo runs where in each run the anchor locations a_i and the source x were randomly generated from a uniform distribution over the square $[-10, 10] \times [-10, 10]$. The observed ranges r_i are generated as described in Section 2, where σ takes on four different values: $1, 10^{-1}, 10^{-2}, 10^{-3}$. Table 1 lists the average squared position errors $\|\hat{x} - x\|^2$ over all realizations, where $\hat{x}$ denotes the solution of SR-LS or SLCP. The best result for each σ is marked in boldface, showing the superiority of SLCP. This example and Table 1 are provided for direct comparison with Example 2 and Table 1 in [3], where SR-LS was shown to outperform other methods.

σ	SR-LS	SLCP
1e-3	3.22e-6	**1.72e-6**
1e-2	1.93e-4	**1.39e-4**
1e-1	2.24e-2	**1.61e-2**
1	2.13	**1.86**

Table 1. Mean squared position error of SR-LS and SLCP

Example 2: To further investigate the accuracy of the methods, the performances of several algorithms (SR-LS, SLCP, SFP initialized by SLCP, SFP initialized by SR-LS) are tested for five anchors with 10 different noise levels, $\sigma \in [0.01, 0.1]$. The performance metric is root mean square error (RMSE), defined as $\sqrt{\frac{1}{K}\sum_{k=1}^{K}\|x - \hat{x}^k\|^2}$, where $\hat{x}^k$ denotes an estimated source position in the k-th Monte Carlo run for the specific noise realization. The number of Monte Carlo runs is $K = 200$. Fig. 3a shows that plain SLCP has better accuracy than SR-LS, although the performance gap closes after iterative refinement by SFP. To compare the RMSE of the algorithms in the presence of outliers, modified range measurements are created according to $r_i = \|x - a_i\| + w_i + |\epsilon|$, where ϵ is a white Gaussian noise term with standard deviation $\sigma_{outlier} \in [0, 1]$. The disturbance $|\epsilon|$ only affects the measured range between the second anchor and the source, whereas w_i with $\sigma = 0.04$ is present in all observations.

Fig. 3b shows an increased RMSE gap between plain SR-LS and SLCP. Regarding convergence speed in SFP refinement, the mean number of iterations, N, over all Monte Carlo runs was calculated for initialization using SLCP or SR-LS. For $\sigma = 0.1$ and no outliers, we have $N_{\text{SLCP+SFP}} = 19$ and $N_{\text{SR-LS+SFP}} = 35$, whereas for $\sigma_{outlier} = 1$ and $\sigma = 0.04$ these become $N_{\text{SLCP+SFP}} = 7$ and $N_{\text{SR-LS+SFP}} = 33$. These

(a) without outlier

(b) with outlier
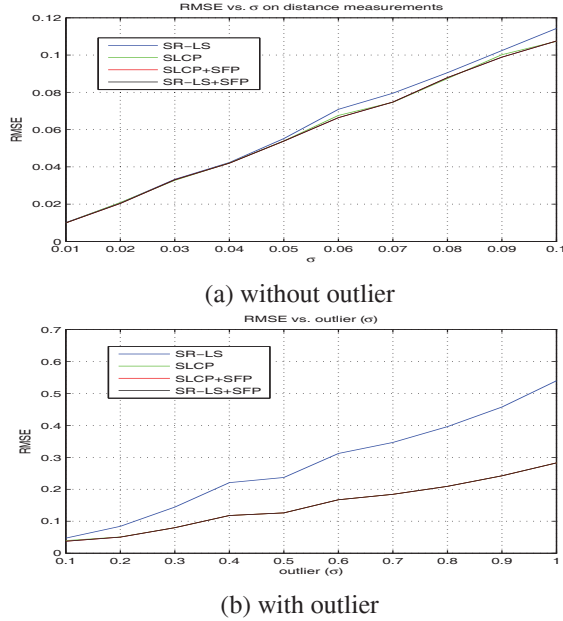
Fig. 3. RMSE vs. disturbance power for five anchors. The curves for SLCP, SR-LS+SFP and SLCP+SFP are nearly overlapping in the two figures.

results show that plain SLCP has better accuracy than plain SR-LS, while their computational times in our implementation are similar. Besides, when the solution of SLCP is given as an initialization to SFP, it significantly reduces the number of iterations compared with SR-LS initialization. Actually, plain SLCP has very good accuracy in this example, and SFP refinement only provides marginal improvements.

The improvement of SLCP over SR-LS is not as significant for three anchors, because the set illustrated in Section 2 more often becomes markedly nonconvex for some noise/anchor realizations, and the SDP relaxation does not yield a desirable nearly rank-1 solution. Extracting the source location by SVD might then be inappropriate.

Example 3: This example demonstrates that, in addition to increasing the convergence speed of SFP, SLCP initialization also alleviates the problem of convergence to local extrema of the ML cost function. The five-anchor setup is similar to that of the previous example for $\sigma \in [0.02, 0.1]$ and no outliers, but the anchor positions are chosen such that the ML cost function has one local minimum in addition to the global one. Fig. 4 shows that now SFP refinement does not close the performance gap between SLCP and SR-LS initialization because in the latter case the algorithm converges more often to the undesirable solution, thus producing a larger RMSE.
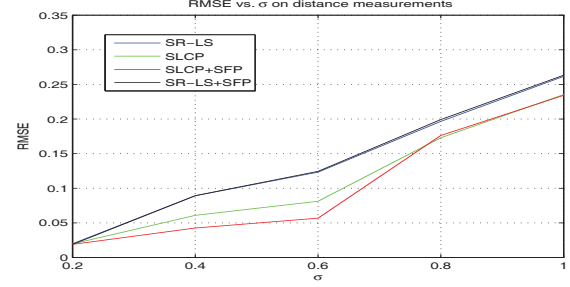
Fig. 4. RMSE vs. disturbance power for five anchors with one local and one global minimum.

4. CONCLUSION

The nonconvex ML based source localization problem was tackled by formulating it as an optimization problem in the complex plane and then taking advantage of the nearly convex nature of the resulting cost function and constraint set to obtain a SDP relaxation. This essentially involves dropping a rank constraint, which was found to have a negligible impact on the accuracy of the solution in many scenarios. Simulation results showed that SLCP provides very accurate results compared to other existing methods. Moreover, when used for initialization of iterative refinement methods it provides a good starting point that reduces both the number of required iterations and the probability of convergence to local extrema.

Future work will focus on finding a better rank-1 approximation than SVD to the matrix output of SLCP when very few anchors are used. Formulating SLCP in higher dimensions is also a topic of interest, e.g., for 3D source localization.

5. REFERENCES

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

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

[3] A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," *IEEE Transactions on Signal Processing*, vol. 56, no. 5, pp. 1770–1778, May 2008.

[4] P. Oğuz-Ekim, J. Gomes, J. Xavier, and P. Oliveira, "ML-based sensor network localization and tracking: Batch and time-recursive approaches," in *Proc. 17th European Signal Processing Conference (EUSIPCO)*, Glasgow, Scotland, August 2009.

[5] M. G. Rabbat and R. D. Nowak, "Decentralized source localization and tracking," in *Proc. ICASSP*, Montreal, Canada, May 2004, vol. 3, pp. 921–924.

[6] A. O. Hero and D. Blatt, "Sensor network source localization via projection onto convex sets (POCS)," in *Proc. ICASSP*, Philadelphia, USA, March 2005, vol. 3, pp. 689–692.

[7] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice-Hall, Englewood Cliffs, NJ, 1993.