An Angular Approach for Range-Based Approximate Maximum Likelihood Source Localization Through Convex Relaxation

Pinar Oğuz-Ekim, *Student Member, IEEE*, João Pedro Gomes, *Member, IEEE*, João Xavier, *Member, IEEE*, Marko Stošić, and Paulo Oliveira, *Senior Member, IEEE*

Abstract—This work considers the problem of locating a single source from noisy range measurements to a set of nodes in a wireless sensor network. We propose two new techniques that we designate as Source Localization with Nuclear Norm (SLNN) and Source Localization with ℓ_1-norm (SL-ℓ_1), which extend to arbitrary real dimensions our prior work on 2D source localization formulated in the complex plane. Our approach is based on formulating a Maximum-Likelihood (ML) estimation problem, and then using convex relaxation techniques to obtain a semidefinite program (SDP) that can be globally and efficiently solved. SLNN directly approximates the Gaussian ML solution, and the relaxation is shown to be tighter than in other methods in the same class. We present an analysis of the convexity properties of the constraint set for the 2D complex version of SLNN (SLCP) to justify the observed tightness of the relaxation. We propose the SL-ℓ_1 algorithm to address the Laplacian noise case, which models the presence of outliers in range measurements. We overcome the non-differentiability of the Laplacian likelihood function by rewriting the ML problem as an exact weighted version of the Gaussian case. In terms of accuracy of localization, the proposed algorithms globally outperform state-of-the-art optimization-based methods in different noise scenarios, while exhibiting moderate computational complexity.

Index Terms—Centralized method, convex hull, convex relaxation, range-based source localization, semidefinite programming.

I. INTRODUCTION

LOCATING a source from range measurements to a set of known reference points (anchors) is a classic problem in many engineering applications (e.g., radar, sonar, GPS, mobile communications), and has received much attention over the years. Recently, source localization from range measurements has been intensively examined in the context of wireless sensor networks (WSN), where ranges estimated from times of arrival, or from surrogates such as received signal strength, are somewhat unreliable due to the complexity of many WSN propagation environments (e.g., indoor settings with few unobstructed line-of-sight (LOS) paths).

Spatial information *per se*, or as georeference to other sensor measurements, is crucial in WSN applications and warrants investigation into suitable localization algorithms. While many approaches to source localization based on classical triangulation or heuristic criteria can be found in the WSN literature [1], [2], our primary focus is on optimization-based methods derived from the likelihood function of observations, or related cost functions [3]–[9]. Through this formal approach we expect to take advantage of the optimality properties of maximum-likelihood (ML) estimates to improve the robustness to perturbations in range measurements.

Centralized ML algorithms for range-based source localization, which require the transmission of the full data set to a fusion node for processing, are proposed in [4]–[12]. Some of these resort to semidefinite relaxation (SDR) to alleviate the problem of algorithmic convergence to undesirable local maxima of the likelihood function under Gaussian noise [4], [9] or Laplacian noise [10]. A related alternative approach proposed in [5] solves a constrained least-squares (LS) problem using squared range (SR) measurements, subject to a quadratic constraint. This was shown to outperform, on average, the ML SDR approach of [9] whose relaxed solutions sometimes fail to produce meaningful source position vectors (rank one solutions). An SDP-based method proposed in [13] is originally developed for sensor network localization and accounts for uncertainties in anchor positions. Another approach, proposed in [7], approximates the ML solution via second-order cone programming and a low-dimensional search. A linearization-based method presented in [12] and [14] solves two consecutive linear problems where the first one is unconstrained and the second is a refinement step that approximately exploits a nonlinear norm constraint linking the optimization variables.

Distributed algorithms for wireless sensor nodes, where the source location is iteratively determined through in-network processing at individual nodes and communication between neighbours, are also being very actively pursued [15]. These techniques, however, are not the focus of our work.

This paper develops algorithms for range-based source localization in arbitrary dimensions, including 2D/3D, through an ML SDR approach. The original principle, proposed in [4]

for 2D localization under Gaussian noise, was termed Source Localization in the Complex Plane (SLCP) and reformulates range-based localization as an equivalent problem of determining a set of angles for the source-anchor directions. The main motivation for the geometrically-inspired approach followed in SLCP is the fact that the reformulated localization problem will often be convex or nearly convex (depending on observations), and thus amenable to convex relaxation, as explained in more detail below. Our relaxation for the nonconvex likelihood function is tighter than the popular one presented in [9], in the sense that the solution will more often have (near) rank-1, as required to obtain target coordinates by factorization. SLCP also outperforms other approaches such as the SR-LS algorithm of [5], which iteratively solves a generalized trust-region subproblem and dispenses with factorization of rank-1 matrices, but undergoes some degradation due to squaring of (noisy) ranges in the cost function. The degradation of SR-LS becomes more severe in the presence of outliers [10], which commonly affect practical range measurement systems, e.g., when non-line-of-sight (NLOS) propagation occurs. The contributions of this work are as follows:

1) The paper expands upon the results of [4], extending the novel angular framework of SLCP from 2D localization, which relied on a formulation where target and anchor coordinates were represented as complex numbers, to arbitrary (real) dimensions. The ability to operate in 3D is especially relevant in indoor localization systems, where the placement of anchors may be dictated by structural constraints of buildings, rather than desirable geometries. We term the new SDR method Source Localization with Nuclear Norm (SLNN), as this norm arises naturally in the cost function of our relaxed optimization problem. Similarly to SLCP, SLNN offers a tight relaxation in most problem instances, and retains a performance advantage over several other methods. Furthermore, we provide an analysis of the geometry and accuracy properties of SLCP, and find that its success in providing tight relaxations relies on certain parametrically defined sets in $\mathbb{R}^2$ being nearly convex.

2) In [10] a modification of SLCP based on the same angular strategy was introduced for ML source localization under Laplacian noise. The new approach, termed SL-ℓ_1, builds robustness to outlier measurements, a property that was observed in simulation even for non-Laplacian range errors. Mathematically, this stems from usage of ℓ_1 instead of ℓ_2 norms in the likelihood function, which tend to de-emphasize large discrepancies between predicted and measured ranges. In this paper, we provide a conceptually similar extension for source localization beyond 2D, consisting of a reformulation of the nondifferentiable log-likelihood function for Laplacian noise as a reweighted version of the Gaussian log-likelihood. We propose SL-ℓ_1 in multiple dimensions (SL-ℓ_1 MD), which outperforms all benchmarked algorithms, and a simplified formulation (SL-ℓ_1 SD), which has slightly worse performance but is less computationally complex.

3) We perform extensive simulations to compare the performance of our algorithms with several others of varying complexity and accuracy based on different criteria. We find that ours exhibit the most consistent performance over different types of range measurement noise, and relatively close to the Cramér-Rao lower bound for Gaussian noise. The complexity is comparable to that of other SDP-based methods, and appears suitable for practical implementation in several centralized scenarios of interest with current technology.

Throughout, scalars, vectors and matrices are denoted by lowercase, boldface lowercase and boldface uppercase letters, respectively. Individual components of matrix $\mathbf{X}$ are written as x_{ij} and those of vector $\mathbf{x}$ as x_i. The superscript $T(H)$ denotes the transpose (Hermitian) of the given real (complex) vector or matrix, $\langle \cdot, \cdot \rangle$ denotes the inner product of two vectors or matrices, and $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. For symmetric matrix $\mathbf{X}$, $\mathbf{X} \succeq 0$ means that $\mathbf{X}$ is positive semidefinite. We denote the Frobenius norm of matrix $\mathbf{X}$ as $\|\mathbf{X}\|_F = \sqrt{\mathrm{tr}(\mathbf{X}^H \mathbf{X})}$ and its nuclear norm as $\|\mathbf{X}\|_N = \mathrm{tr}((\mathbf{X}^H \mathbf{X})^{1/2})$. $\|\mathbf{x}\|$ corresponds to the Euclidean norm of vector $\mathbf{x}$. Below, $\mathbf{I}_m$ is the $m \times m$ identity matrix and $\mathbf{1}_m$ is the vector of m ones. The convex hull of set $\mathcal{S}$ is denoted by $\mathrm{co}(\mathcal{S})$.

The paper is organized as follows. In Section II we formulate the ML source location problem under Gaussian or Laplacian noise. In Section III we derive the SLCP algorithm for 2D localization, analyzing the geometry of the associated optimization problem and the tightness of the relaxation in Appendix A. In Section IV we derive the SLNN algorithm (Section IV-A) and SL-ℓ_1MD/SL-ℓ_1 SD (Section IV-B), which extend SLCP and SL-ℓ_1 to 3 (and higher) dimensions, respectively. Section V benchmarks our algorithms against performance bounds and other available algorithms in different representative classes using simulated data. Finally, conclusions are drawn in Section VI.

II. PROBLEM FORMULATION

Let $\mathbf{x} \in \mathbb{R}^n$ be the unknown source position, $\mathbf{a}_i \in \mathbb{R}^n$, $i = 1, \ldots, m$ be known sensor positions (anchors), and $r_i = \|\mathbf{x} - \mathbf{a}_i\| + w_i$ be the measured range between the source and the ith anchor, where w_i denotes a noise term with standard deviation σ. Under independent and identically distributed (i.i.d.) Gaussian or Laplacian noise, maximizing the likelihood of observations for the source localization problem is equivalent to

$$\underset{\mathbf{x}}{\text{minimize}} \sum_{i=1}^{m} |\|\mathbf{x} - \mathbf{a}_i\|^p - r_i^p|^q \tag{1}$$

for suitable values of p and q. We will derive the SLCP/SLNN algorithms to (approximately) solve (1) under Gaussian noise ($p = 1$, $q = 2$), whereas SL-ℓ_1 will solve it under Laplacian noise ($p = 1$, $q = 1$). The case ($p = 2$, $q = 2$) is also of interest and corresponds to the cost function used in the SR-LS algorithm of [5], which is used to benchmark our algorithms. Note that the cost function for SR-LS is not a likelihood function, and it arises out of mathematical convenience, at the cost of increasing the sensitivity to measurement noise due to the presence of third and fourth powers of r_i.
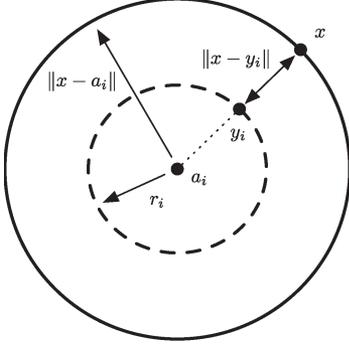
Fig. 1. Geometrical interpretation of terms in the cost function (1) for $p = 1$, $q = 2$.

The main difficulties of solving (1) lie in the fact that this cost function is, in general, nonconvex and possibly multimodal. We address the nonconvexity of the cost function in Sections III and IV by developing convex relaxations that turn out to be tight in most problem instances, thus providing a very good approximation to the global minimum of (1). An additional challenge is posed by the nondifferentiability of (1) for $q = 1$, which is addressed in Section IV-B by rewriting it as a weighted version of the case $q = 2$, where the weights themselves become optimization variables.

III. SOURCE LOCALIZATION IN 2D: SLCP

For $p = 1$, $q = 2$, each term in (1) can be viewed as the squared distance between two circles centered on $\mathbf{a}_i$, one with radius $\|\mathbf{x} - \mathbf{a}_i\|$, and the other with radius r_i (see Fig. 1). This term can be replaced by the squared norm of the difference between the position vector $\mathbf{x}$ and its closest point on the circle $\{\mathbf{y} \in \mathbb{R}^2 : \|\mathbf{y} - \mathbf{a}_i\| = r_i\}$, which we denote by $\mathbf{y}_i$. Problem (1) can then be equivalently expressed as (a formal proof of equivalence is provided in [10])

$$\underset{\mathbf{x}, \mathbf{y}_i}{\text{minimize}} \quad \sum_{i=1}^{m} \|\mathbf{x} - \mathbf{y}_i\|^2$$
$$\text{subject to} \quad \|\mathbf{y}_i - \mathbf{a}_i\| = r_i \, i = 1, \ldots, m. \quad (2)$$

If $\mathbf{y}_i$ is fixed, the solution of (2) with respect to $\mathbf{x}$ is an unconstrained optimization problem whose solution is readily obtained as the center of mass of the constellation $\mathbf{x} = (1/m) \sum_{i=1}^{m} \mathbf{y}_i$. Moreover, in 2D the constraints of (2) can be compactly described in the complex plane, yielding

$$\underset{\mathbf{y}, \boldsymbol{\theta}}{\text{minimize}} \quad \left\| \frac{1}{m} \mathbf{1}_m \mathbf{1}_m^T \mathbf{y} - \mathbf{y} \right\|^2$$
$$\text{subject to} \quad \mathbf{y} = \mathbf{a} + R\boldsymbol{\theta}, \quad (3)$$

where $\mathbf{a} = [a_1 \ldots a_m]^T \in \mathbb{C}^m$ holds the anchor coordinates, expressed as complex numbers, $\mathbf{R} = \text{diag}(r_1, \ldots, r_m) \in \mathbb{R}^{m \times m}$, and $\boldsymbol{\theta} = [e^{j\phi_1} \ldots e^{j\phi_m}]^T \in \mathbb{C}^m$. The problem acquires a flavor of angle-of-arrival localization, as the angles ϕ_i are variables that encode a set of directions, departing from anchor nodes, that ideally intersect at the source position. The complex representation makes it simple to impose unit magnitude constraints on the elements of $\boldsymbol{\theta}$, and later relax them to obtain an

SDR. Expanding the cost function and deleting constant terms yields a quadratic constrained problem

$$\underset{\boldsymbol{\theta}}{\text{minimize}} \quad 2\text{Re}(\mathbf{c}^H \boldsymbol{\theta}) - \frac{1}{m} \boldsymbol{\theta}^H \mathbf{r} \mathbf{r}^T \boldsymbol{\theta}$$
$$\text{subject to} \quad |\theta_i| = 1, \quad (4)$$

where $\mathbf{r} = \mathbf{R}\mathbf{1}_m$ and $\mathbf{c} = \mathbf{R}(\mathbf{I}_m - (1/m)\mathbf{1}_m\mathbf{1}_m^T)\mathbf{a}$.

To proceed we now wish to replace $\text{Re}(\mathbf{c}^H \boldsymbol{\theta})$ in (4) with $-|\mathbf{c}^H \boldsymbol{\theta}|$, which is readily written as a function of a quadratic form in $\boldsymbol{\theta}$ and then relaxed in the same way as the second term in the objective function. To this end, first note that if $\boldsymbol{\theta}$ is replaced with $\boldsymbol{\theta} e^{j\gamma}$ neither the second term in the objective function of (4) nor the constraints change for any angle γ. By proper choice of γ the complex number $\mathbf{c}^H \boldsymbol{\theta}$ may be rotated to the (negative) real axis for any feasible $\boldsymbol{\theta}$, such that $\text{Re}(\mathbf{c}^H \boldsymbol{\theta} e^{j\gamma}) = -|\mathbf{c}^H \boldsymbol{\theta}|$, thus reducing the value of the objective function relative to other values of γ. This implies that any optimal solution of (4) will satisfy $\text{Re}(\mathbf{c}^H \boldsymbol{\theta}) = -|\mathbf{c}^H \boldsymbol{\theta}|$, which justifies replacing $\text{Re}(\cdot)$ with $-|\cdot|$ in the cost function. It should be kept in mind, however, that once a solution $\boldsymbol{\theta}$ to the modified optimization problem is obtained it should be rotated to obtain the actual vector of phases $\boldsymbol{\theta} e^{j\gamma}$ such that $\text{Re}(\mathbf{c}^H \boldsymbol{\theta} e^{j\gamma}) = -|\mathbf{c}^H \boldsymbol{\theta}|$.

Now the modified problem is equivalently written as

$$\underset{\boldsymbol{\theta}}{\text{maximize}} \quad 2\sqrt{\text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\theta}\boldsymbol{\theta}^H)} + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\theta}\boldsymbol{\theta}^H)$$
$$\text{subject to} \quad |\theta_i| = 1, \quad (5)$$

and following standard manipulations we introduce the new variable $\boldsymbol{\Phi} = \boldsymbol{\theta}\boldsymbol{\theta}^H$ and an associated (nonconvex) constraint $\text{rank}(\boldsymbol{\Phi}) = 1$. Finally, a SDR formulation of SLCP is obtained by introducing the hypograph variable t such that $0 \leq t \leq 2\sqrt{\text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\Phi})}$ and dropping the rank constraint

$$\underset{\boldsymbol{\Phi}, t}{\text{maximize}} \quad t + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\Phi})$$
$$\text{subject to} \quad \boldsymbol{\Phi} \succeq 0, \, \phi_{ii} = 1, \quad 4\mathbf{c}^H \boldsymbol{\Phi}\mathbf{c} \geq t^2. \quad (6)$$

Remark that the solution of (6) is a positive semidefinite matrix, which should have a clearly dominant eigenvalue in problem instances where the SDR is an accurate approximation to the initial problem (2). In such cases $\boldsymbol{\Phi} \approx \lambda_1 \mathbf{u}_1 \mathbf{u}_1^H$, where λ_1 is the highest eigenvalue of $\boldsymbol{\Phi}$ and $\mathbf{u}_1$ the corresponding eigenvector, and the vector of complex phases is estimated as $\boldsymbol{\theta} = \sqrt{\lambda_1} \mathbf{u}_1$. In Appendix A we examine the geometry of the cost function and constraint set in (5), shedding light on the observed high probability of obtaining a nearly rank-1 solution.

IV. SOURCE LOCALIZATION IN HIGHER DIMENSIONS

In this section, we derive extensions of the localization algorithms in [4] and in [10] for 3D and higher dimensions under both Gaussian and Laplacian noise assumptions. In the first part we develop the SLNN algorithm for Gaussian noise. In Section IV-B we then present the SL-ℓ_1 MD and SL-ℓ_1 SD algorithms for Laplacian noise, which overcome nondifferentiability issues by expressing the likelihood function as a variable-weighted Gaussian likelihood.

A. Localization Under Gaussian Noise: SLNN

To extend the approach used in SLCP to $n > 2$ dimensions, we write the circle/sphere equations in (2) using an equivalent parametric form with real coordinates

$$\begin{array}{cl} \underset{\mathbf{x}, \mathbf{y}_i, \mathbf{u}_i}{\text{maximize}} & \sum_{i=1}^{m} \|\mathbf{x} - \mathbf{y}_i\|^2 \\ \text{subject to} & \mathbf{y}_i = \mathbf{a}_i + r_i \mathbf{u}_i, \quad \|\mathbf{u}_i\| = 1, \end{array} \quad (7)$$

where $\mathbf{x}, \mathbf{y}_i, \mathbf{a}_i$, and $\mathbf{u}_i$ are now vectors in $\mathbb{R}^n$, rather than complex scalars used in SLCP. In (7) variable $\mathbf{u}_i \in \mathbb{R}^n$ is a unit-norm vector that plays the same role as the complex phase shift $e^{j\phi_i}$ in SLCP. The derivation parallels that given for SLCP, but dealing with vectors for the angular components leads to significant differences, most notably a rotation subproblem whose solution is no longer obvious and turns out to be given by the nuclear norm of the linearly transformed angular matrix. We write (7) equivalently as

$$\begin{array}{cl} \underset{\mathbf{x}, \mathbf{y}_i, \mathbf{u}_i}{\text{maximize}} & \|\mathbf{1}_m \mathbf{x}^T - \mathbf{Y}\|_F^2 \\[2mm] \text{subject to} & \underbrace{\begin{bmatrix} \mathbf{y}_1^T \\ \vdots \\ \mathbf{y}_m^T \end{bmatrix}}_{\mathbf{Y}} = \underbrace{\begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}}_{\mathbf{A}} + \mathbf{R} \underbrace{\begin{bmatrix} \mathbf{u}_1^T \\ \vdots \\ \mathbf{u}_m^T \end{bmatrix}}_{\mathbf{U}}, \quad \|\mathbf{u}_i\| = 1, \end{array} \quad (8)$$

where $\mathbf{R} = \text{diag}(r_1, \ldots, r_m)$ as in (3). For fixed $\mathbf{y}_i$, $\mathbf{u}_i$ (8) describes n uncoupled least-squares problems whose variables are the components of the source location vector $\mathbf{x}$. The optimal solutions may be jointly written compactly as

$$\mathbf{x}^T = \left(\mathbf{1}_m^T \mathbf{1}_m\right)^{-1} \mathbf{1}_m^T \mathbf{Y} = \frac{1}{m} \mathbf{1}_m^T \mathbf{Y}. \quad (9)$$

After replacing $\mathbf{x}$ in (8) with (9), the objective function becomes $\|\mathbf{\Pi}\mathbf{Y}\|_F^2 = \text{tr}(\mathbf{Y}^T \mathbf{\Pi} \mathbf{Y})$, where $\mathbf{\Pi} = \mathbf{I}_m - (1/m)\mathbf{1}_m \mathbf{1}_m^T$ is a projection matrix (hence idempotent). Similarly to (3), and (4), we can now eliminate variable $\mathbf{Y}$ and the first set of equality constraints, expand its definition in the objective function, and ignore constant terms to obtain

$$\begin{array}{cl} \underset{\mathbf{U}}{\text{maximize}} & 2\,\text{tr}(\mathbf{C}^T \mathbf{U}) - \frac{1}{m}\text{tr}(\mathbf{U}^T \mathbf{r}\mathbf{r}^T \mathbf{U}) \\ \text{subject to} & \|\mathbf{u}_i\| = 1, \end{array} \quad (10)$$

where $\mathbf{C} = \mathbf{R}\mathbf{\Pi}\mathbf{A}$ and, as in (4), $\mathbf{r} = \mathbf{R}\mathbf{1}_m$.

Nuclear Norm Approximation: As in the complex formulation we wish to rewrite the first term in the objective function of (10) in a form that is more amenable to SDR. Similarly to SLCP, where we introduced a complex rotation $e^{j\gamma}$ to progress from (4) to (5), in the new optimization problem $\mathbf{U}$ is replaced with the product $\mathbf{U}\mathbf{V}$, where $\mathbf{V}$ is an $n \times n$ orthogonal matrix $(\mathbf{V}^T\mathbf{V} = \mathbf{V}\mathbf{V}^T = \mathbf{I}_n)$ that plays the role of a rotation factor. This yields

$$\begin{array}{cl} \underset{\mathbf{U}, \mathbf{V}}{\text{maximize}} & 2\,\text{tr}(\mathbf{C}^T \mathbf{U}\mathbf{V}) - \frac{1}{m}\text{tr}(\mathbf{V}^T \mathbf{U}^T \mathbf{r}\mathbf{r}^T \mathbf{U}\mathbf{V}) \\ \text{subject to} & \|\mathbf{u}_i\| = 1, \quad \mathbf{V}^T\mathbf{V} = \mathbf{I}_n. \end{array} \quad (11)$$

Note that, due to the orthogonality of $\mathbf{V}$, each row of $\mathbf{U}\mathbf{V}$ still has unit norm, so for any feasible $\mathbf{U}$ in (10) $\mathbf{U}\mathbf{V}$ is also feasible. Regarding (11), $\mathbf{V}$ may be interpreted as an inner optimization variable that, for each candidate $\mathbf{U}$, minimizes the value of the objective function. Noting that the second term in the objective function (11) does not depend on $\mathbf{V}$, as $\text{tr}(\mathbf{V}^T\mathbf{U}^T\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{V}) = \text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{V}\mathbf{V}^T\mathbf{U}^T) = \text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{U}^T)$, the inner optimization problem simply becomes

$$\begin{array}{cl} \underset{\mathbf{V}}{\text{maximize}} & \text{tr}(\mathbf{C}^T \mathbf{U}\mathbf{V}) = \langle \mathbf{V}, \mathbf{U}^T \mathbf{C} \rangle \\ \text{subject to} & \mathbf{V}^T\mathbf{V} = \mathbf{I}_n. \end{array} \quad (12)$$

This involves the minimization of a linear function on the set of orthogonal matrices, which resembles the known problem of minimizing a linear function of a vector $\mathbf{v}$, say, $\langle \mathbf{v}, \mathbf{a} \rangle$, on the unit sphere $\|\mathbf{v}\|^2 = \mathbf{v}^T\mathbf{v} = 1$. The minimizer for this problem is readily seen to be the point on the sphere along vector $-\mathbf{a}$, with optimal cost $-\|\mathbf{a}\|$ [16]. With $\mathbf{U}^T\mathbf{C}$ playing the role of $\mathbf{a}$, one would therefore expect the optimal value of (12) to equal $-\|\mathbf{C}^T\mathbf{U}\|$, for some matrix norm of $\mathbf{C}^T\mathbf{U}$. In Appendix C it is shown that this is indeed the case, and the appropriate norm to consider turns out to be the *nuclear norm*, defined for matrix $\mathbf{X}$ as $\|\mathbf{X}\|_N = \text{tr}((\mathbf{X}^H\mathbf{X})^{1/2})$ and equal to the sum of its singular values. The optimization problem (11) is therefore equivalently rewritten as

$$\begin{array}{cl} \underset{\mathbf{U}}{\text{maximize}} & -2\|\mathbf{C}^T\mathbf{U}\|_N - \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{U}^T) \\ \text{subject to} & \|\mathbf{u}_i\| = 1, \end{array} \quad (13)$$

or

$$\begin{array}{cl} \underset{\mathbf{U}}{\text{maximize}} & 2\,\text{tr}\left((\mathbf{C}^T\mathbf{U}\mathbf{U}^T\mathbf{C})^{\frac{1}{2}}\right) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{U}\mathbf{U}^T) \\ \text{subject to} & \|\mathbf{u}_i\| = 1. \end{array} \quad (14)$$

We now introduce the variable $\mathbf{W} = \mathbf{U}\mathbf{U}^T$ and ignore the associated nonconvex constraint $\text{rank}(\mathbf{W}) = n$ to obtain the SDR

$$\begin{array}{cl} \underset{\mathbf{W}}{\text{maximize}} & 2\,\text{tr}\left((\mathbf{C}^T\mathbf{W}\mathbf{C})^{\frac{1}{2}}\right) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{W}) \\ \text{subject to} & \mathbf{W} \succeq 0, \quad w_{ii} = 1. \end{array} \quad (15)$$

The objective function of (15) is the sum of a concave[1] function of $\mathbf{W}$ with a linear term, and is therefore concave. The constraint set of (15) is convex, thus establishing that this is indeed a convex optimization problem which is expressed in standard SDP form as

$$\begin{array}{cl} \underset{\mathbf{W}, \mathbf{Z}}{\text{maximize}} & 2\,\text{tr}(\mathbf{Z}) + \frac{1}{m}\text{tr}(\mathbf{r}\mathbf{r}^T\mathbf{W}) \\[2mm] \text{subject to} & \mathbf{W} \succeq 0, \ w_{ii} = 1, \ \mathbf{Z} \succeq 0, \ \begin{bmatrix} \mathbf{C}^T\mathbf{W}\mathbf{C} & \mathbf{Z} \\ \mathbf{Z} & \mathbf{I}_n \end{bmatrix} \succeq 0. \end{array} \quad (16)$$

The equivalence between (15) and (16) is proved in Appendix C.

---

[1] The first term is the composition of the linear map $\mathbf{X} = \mathbf{C}^T\mathbf{W}\mathbf{C}$ with $\text{tr}(\mathbf{X}^{1/2})$, which is known to be concave in $\mathbf{X}$ [16].

Similar to the complex 2D formulation, the solution of our SDR is a $m \times m$ matrix $\mathbf{W}$ that should have approximately rank n when the relaxation is tight. The matrix $\mathbf{U}$ of unit-norm vectors is obtained by Singular Value Decomposition (SVD) factorization of $\mathbf{W}$ and, after accounting for the inner rotation of $\mathbf{U}$, it is used to build the $\mathbf{y}_i$ and, ultimately, the source position vector $\mathbf{x}$. Algorithm 1 summarizes the steps of SLNN.

Algorithm 1 Summary of SLNN

1: Given the anchor positions and range measurements, solve the SDR (16)
2: Compute a rank-n approximation of the SDR solution as $\mathbf{W} \approx \mathbf{U}\mathbf{U}^T$
3: Solve the inner optimization problem (12) to get the rotation matrix $\mathbf{V}$ (Appendix C)
4: Obtain the matrix of sphere projections as $\mathbf{Y} = \mathbf{A} + \mathbf{RUV}$
5: Estimate the source position as the centroid of the rows of $\mathbf{Y}$, $\mathbf{x} = (1/m)\mathbf{Y}^T\mathbf{1}_m$

B. Localization Under Laplacian Noise: SL-ℓ_1

When disturbances are Laplacian and i.i.d., thus heavier tailed than Gaussian, maximizing the likelihood amounts to solving (1) for $p = q = 1$

$$\underset{\mathbf{x}}{\text{minimize}} \quad \sum_{i=1}^{m} |\|\mathbf{x} - \mathbf{a}_i\| - r_i|. \tag{17}$$

The presence of $|\cdot|$ in each summation term of (17), rather than $(\cdot)^2$, de-emphasizes the contributions of measurements r_i corrupted by large noise values. The optimal point of (17) is thus less biased by these outlier measurements than the cost function (1) for the Gaussian case $p = 1$, $q = 2$. However, a major difficulty in solving (17) is the fact that the cost function is not differentiable, making it less amenable to the types of analytic manipulations that we use to develop SDR. The strategy that we adopt to circumvent this difficulty parallels the one used in [10] for 2D sources, and as a key ingredient involves squaring the cost function of (17) (which does not affect the location of extremal points), and then rewriting it as

$$\underset{\mathbf{x}, \boldsymbol{\lambda}}{\text{minimize}} \quad \sum_{i=1}^{m} \frac{(\|\mathbf{x} - \mathbf{a}_i\| - r_i)^2}{\lambda_i}$$

$$\text{subject to} \quad \lambda_i > 0, \quad \mathbf{1}_m^T\boldsymbol{\lambda} = 1. \tag{18}$$

The cost function is thus reduced to a weighted version of the more tractable Gaussian log-likelihood, where the real weighting coefficients λ_i become optimization variables themselves. See [10] for a proof of this result (also [17]). Salient differences of the derivation to that for the complex case [10] include the presence of nonconvex rank-1 constraints in the reformulation based on epigraph variables of Section IV-B1, which are approximately enforced through a penalizing term in the cost function. The internal structure of variables is also different. Now, the manipulations used earlier in Section IV-A for the

development of SLNN can be replicated here to reformulate the problem as

$$\underset{\mathbf{x}, \mathbf{y}_i, \mathbf{u}_i, \boldsymbol{\lambda}}{\text{minimize}} \quad \sum_{i=1}^{m} \frac{\|\mathbf{x} - \mathbf{y}_i\|^2}{\lambda_i}$$

$$\text{subject to} \quad \mathbf{y}_i = \mathbf{a}_i + r_i\mathbf{u}_i, \quad \|\mathbf{u}_i\| = 1, \lambda_i > 0, \mathbf{1}_m^T\boldsymbol{\lambda} = 1. \tag{19}$$

For given $\mathbf{y}_i$, $\mathbf{u}_i$, and $\boldsymbol{\lambda}$, (19) has a least-squares cost function whose unconstrained optimal solution with respect to $\mathbf{x}$ is readily found in closed form as $\mathbf{x}^* = (\sum_{i=1}^{m} \lambda_i^{-1})^{-1} \sum_{i=1}^{m} \lambda_i^{-1}\mathbf{y}_i$. Substituting the optimal $\mathbf{x}$ in (19), and using matrix notation, the cost function becomes $\text{tr}(\mathbf{Y}^T\boldsymbol{\Xi}\mathbf{Y})$, where $\boldsymbol{\Xi}$ is the modified projector

$$\boldsymbol{\Xi} = \begin{bmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_m} \end{bmatrix} - \frac{1}{\sum_{i=1}^{m} \frac{1}{\lambda_i}} \begin{bmatrix} \frac{1}{\lambda_1} \\ \vdots \\ \frac{1}{\lambda_m} \end{bmatrix} \begin{bmatrix} \frac{1}{\lambda_1} & \cdots & \frac{1}{\lambda_m} \end{bmatrix}$$

$$= \boldsymbol{\Lambda}^{-1} - \boldsymbol{\Lambda}^{-1}\mathbf{1}(\mathbf{1}^T\boldsymbol{\Lambda}^{-1}\mathbf{1})^{-1}\mathbf{1}^T\boldsymbol{\Lambda}^{-1} \tag{20}$$

with $\boldsymbol{\lambda} = \text{diag}(\lambda_1, \ldots, \lambda_m)$. Now $\mathbf{Y}$ and the related constraint can be eliminated to obtain

$$\underset{\mathbf{U}, \boldsymbol{\lambda}}{\text{minimize}} \quad \text{tr}\left((\mathbf{A} + \mathbf{RU})^T\boldsymbol{\Xi}(\mathbf{A} + \mathbf{RU})\right)$$

$$\text{subject to} \quad \|\mathbf{u}_i\| = 1, \lambda_i > 0, \quad \mathbf{1}_m^T\boldsymbol{\lambda} = 1. \tag{21}$$

1) SL-ℓ_1 in Multiple Dimensions (SL-ℓ_1 MD): An epigraph variable t_i is introduced for each term contributing to $\text{tr}(\cdot)$ in the cost function of (21)

$$\underset{\mathbf{U}, \boldsymbol{\lambda}, \mathbf{t}}{\text{minimize}} \quad \mathbf{t}\mathbf{1}_n$$

$$\text{subject to} \quad \mathbf{e}_i^T(\mathbf{A} + \mathbf{RU})^T\boldsymbol{\Xi}(\mathbf{A} + \mathbf{RU})\mathbf{e}_i \leq t_i,$$

$$\|\mathbf{u}_i\| = 1, \quad \lambda_i > 0, \quad \mathbf{1}_m^T\boldsymbol{\lambda} = 1, \tag{22}$$

where $\mathbf{t} = [t_1 \ldots t_n]$ and $\mathbf{e}_i$ is the standard coordinate vector with 1 in the ith position and zeros elsewhere. As in [10] the matrix inversion lemma is invoked to express (20) as the limiting case of (positive semidefinite) $\boldsymbol{\Xi} = \lim_{\eta \to \infty} (\boldsymbol{\lambda} + \eta\mathbf{1}_m\mathbf{1}_m^T)^{-1}$, which is more amenable to analytic manipulations in optimization problems. In practice we take η as a sufficiently large constant. Using Schur complements the inequality constraint in (22) may be successively written as

$$\begin{bmatrix} t_i & \mathbf{e}_i^T(\mathbf{A} + \mathbf{RU})^T \\ (\mathbf{A} + \mathbf{RU})\mathbf{e}_i & \boldsymbol{\Xi}^{-1} \end{bmatrix} \succeq 0 \tag{23}$$

$$t_i(\boldsymbol{\Lambda} + \eta\mathbf{1}_m\mathbf{1}_m^T) - (\mathbf{A} + \mathbf{RU})\mathbf{e}_i\mathbf{e}_i^T(\mathbf{A} + \mathbf{RU})^T \succeq 0. \tag{24}$$

The last inequality is bilinear in t_i and $\lambda_1, \ldots, \lambda_m$, and we linearize it by replacing the optimization variable $\boldsymbol{\lambda}$ with a new $\boldsymbol{\beta}_i = t_i\boldsymbol{\lambda}$. Now, the $\boldsymbol{\beta}_i$ can be assembled into a matrix

$$\boldsymbol{\beta} = [\boldsymbol{\beta}_1 \; \cdots \; \boldsymbol{\beta}_n] = \boldsymbol{\lambda}\mathbf{t} \tag{25}$$

which, as shown above, should have rank 1 and satisfy $\beta_{ij} > 0$, $\mathbf{1}_m^T\boldsymbol{\beta} = \mathbf{t}$. However, the rank-1 constraint for $\boldsymbol{\beta}$ cannot be

directly imposed in convex formulations, and we resort to a common technique to indirectly induce low rank in optimal solutions by adding to the cost function the (scaled) nuclear norm $\|\boldsymbol{\beta}\|_N$.

Regarding the second term on the left-hand side of (24), we first note that

$$(\mathbf{A} + \mathbf{RU})\mathbf{e}_i = [\mathbf{Ae}_i\ \mathbf{R}]\begin{bmatrix} 1 \\ \mathbf{Ue}_i \end{bmatrix} = [\boldsymbol{\alpha}_i\ \mathbf{R}]\begin{bmatrix} 1 \\ \boldsymbol{v}_i \end{bmatrix} \qquad (26)$$

where $\boldsymbol{\alpha}_i$ and $\boldsymbol{v}_i$ denote the ith columns of matrices $\mathbf{A}$ and $\mathbf{U}$, respectively. Now, consider the following variable, obtained from the stacked rotation vectors that make up $\mathbf{U}$

$$\mathbf{W} = \begin{bmatrix} 1 \\ \text{vec}(\mathbf{U}^T) \end{bmatrix} \begin{bmatrix} 1 & \text{vec}(\mathbf{U}^T)^T \end{bmatrix}$$

$$= \begin{bmatrix} 1 & \mathbf{u}_1^T & \cdots & \mathbf{u}_m^T \\ \mathbf{u}_1 & \underbrace{\mathbf{u}_1\mathbf{u}_1^T}_{\mathbf{W}_{11}} & & \\ \vdots & & \ddots & \\ \mathbf{u}_m & & & \underbrace{\mathbf{u}_m\mathbf{u}_m^T}_{\mathbf{W}_{mm}} \end{bmatrix}. \qquad (27)$$

Further, let I_i denote the set of row indices that extracts the elements of $[1\boldsymbol{v}_i^T]^T$ in (26) from the first column of $\mathbf{W}$. Then, the dyad $\mathbf{W}_{I_iI_i} = \begin{bmatrix} 1 \\ \boldsymbol{v}_i \end{bmatrix}[1\boldsymbol{v}_i^T]$ is readily obtained by selecting the submatrix formed from the I_i rows and I_i columns of $\mathbf{W}$ and this carries over to (24) through (26), which can therefore be written in terms of submatrix $\mathbf{W}_{I_iI_i}$. The positive semidefinite matrix $\mathbf{W}$ will replace $\mathbf{U}$ as an optimization variable, retaining the constraints along the diagonal blocks in (27), namely, $\text{tr}(\mathbf{W}_{ii}) = 1$. Finally, we obtain the convex relaxation of (22) by combining all the above elements and dropping the rank-1 constraint for $\mathbf{W}$ that is implied by (27)

$$\underset{\mathbf{W},\boldsymbol{\beta},\mathbf{t}}{\text{minimize}} \quad \mathbf{t}\mathbf{1}_n + \mu\|\boldsymbol{\beta}\|_N$$

$$\text{subject to} \quad \text{diag}(\boldsymbol{\beta}_i) + t_i\eta\mathbf{1}_m\mathbf{1}_m^T \succeq [\boldsymbol{\alpha}_i\ \mathbf{R}]\mathbf{W}_{I_iI_i}\begin{bmatrix} \boldsymbol{\alpha}_i^T \\ \mathbf{R} \end{bmatrix}$$

$$\mathbf{W} \succeq 0, w_{11} = 1, \text{tr}(\mathbf{W}_{ii}) = 1, \beta_{ij} > 0, \mathbf{1}_m^T\boldsymbol{\beta} = \mathbf{t}. \qquad (28)$$

In the numerical assessment of Section V it was observed that the performance of this algorithm is only very weakly affected by the choice of penalization factor μ over a broad interval, and fixing it *a priori* seems acceptable ($\mu = 10$ was used).

2) Simplified Formulation (SL-ℓ_1 SD): Our simulation results suggest that in most scenarios the accuracy of the solution obtained from (28) is nearly identical to that of a simplified formulation where a single epigraph variable, t, is used. Referring to (22), we now minimize $\text{tr}(t\mathbf{I}_n)$ or, equivalently, t, and replace the first constraint for all $i = 1, \ldots, n$ with the single matrix inequality $(\mathbf{A} + \mathbf{RU})^T\boldsymbol{\Xi}(\mathbf{A} + \mathbf{RU}) \preceq t\mathbf{I}_n$. Applying Schur complements as in (23) and (24) yields

$$t(\boldsymbol{\Lambda} + \eta\mathbf{1}_m\mathbf{1}_m^T) - [\mathbf{A}\ \mathbf{R}]\begin{bmatrix} \mathbf{I}_n \\ \mathbf{U} \end{bmatrix}[\mathbf{I}_n\ \mathbf{U}^T]\begin{bmatrix} \mathbf{A}^T \\ \mathbf{R} \end{bmatrix} \succeq 0 \qquad (29)$$

and again variable $\boldsymbol{\lambda}$ is replaced with $\boldsymbol{\beta} = t\boldsymbol{\lambda}$ such that $\beta_i > 0$, $\mathbf{1}_m^T\boldsymbol{\beta} = t$. Now, however, there is no need to assemble a matrix as in (25) and to include its nuclear norm as a penalization term in the cost function. Finally, to obtain a convex relaxation we replace $\mathbf{U}$ with the new variable $\mathbf{W} = \begin{bmatrix} \mathbf{I}_n \\ \mathbf{U} \end{bmatrix}[\mathbf{I}_n\ \mathbf{U}^T] = \begin{bmatrix} \underbrace{\mathbf{I}_n}_{\mathbf{W}_{11}} & \mathbf{U}^T \\ \mathbf{U} & \mathbf{U}\mathbf{U}^T \end{bmatrix}$, and drop the rank-$n$ constraint on $\mathbf{W}$. This simplified SDP formulation, denoted by SL-ℓ_1 SD, is given by

$$\underset{\mathbf{W},\boldsymbol{\beta},t}{\text{minimize}} \quad t$$

$$\text{subject to} \quad \text{diag}(\boldsymbol{\beta}) + t\eta\mathbf{1}_m\mathbf{1}_m^T \succeq [\mathbf{A}\ \mathbf{R}]\mathbf{W}\begin{bmatrix} \mathbf{A}^T \\ \mathbf{R} \end{bmatrix},$$

$$\mathbf{W} \succeq 0, \mathbf{W}_{11} = \mathbf{I}_n, w_{ii} = 1, \beta_i > 0, \mathbf{1}_m^T\boldsymbol{\beta} = t. \qquad (30)$$

Note that the optimization variables $\mathbf{W}$ and $\boldsymbol{\beta}$ in (28) have size $(mn + 1) \times (mn + 1)$ and $m \times n$, respectively, whereas the corresponding sizes in (30) are only $(m + n) \times (m + n)$ and $m \times 1$. For ambient dimension $n = 2$ or 3 and for $m \approx 5$ anchors used in our simulations, problem (30) has considerably fewer variables than (28), and the gap increases as m and n grow.

Given the configuration for variable $\mathbf{W}$ in both formulations of SL-ℓ_1, the required elements of the rotation vectors that make up $\mathbf{U}$ can be obtained from the rightmost (block) column of $\mathbf{W}$ or by factorizing submatrices along the block diagonal. The former approach is usually more accurate [18]. Algorithm 2 summarizes the steps of SL-ℓ_1 MD and SL-ℓ_1 SD.

Algorithm 2 Summary of SL-ℓ_1 MD/SD

1: Given the anchor positions and range measurements, solve the SDR (28) or (30)
2: Compute a rank-n approximation of the SDR solution from the rightmost (block) column of $\mathbf{W}$ or by factorizing submatrices along the block diagonal
3: Obtain the matrix of sphere projections as $\mathbf{Y} = \mathbf{A} + \mathbf{RU}$
4: Estimate the source position as $\mathbf{x} = (\sum_{i=1}^m(\mathbf{y}_i/\lambda_i)/\sum_{i=1}^m(1/\lambda_i))$

V. PERFORMANCE ANALYSIS AND NUMERICAL RESULTS

In this section, we assess the performance and computational complexity of several localization algorithms. Our goal is to demonstrate that the proposed algorithms offer superior accuracy and consistent performance for various noise assumptions, while retaining a moderate computational load that is compatible with the resources of many practical centralized setups. In Table I, the acronyms of the proposed and benchmarked algorithms used in the figures and their descriptions are presented. ML location estimates should be regarded as lower performance bounds, as the iterative minimization algorithms are initialized with true source positions.

TABLE I
ALGORITHMS USED FOR PERFORMANCE ASSESSMENT IN SIMULATIONS

Algorithms	Description
SLCP	a semidefinite relaxation based method which uses complex plane for Gaussian noise, Section III, [4]
SL-ℓ_1	SLCP's complex counterpart for Laplacian noise, [10]
SLNN	SLCP's higher dimension extension, Section IV-A, Algorithm 1
SL-ℓ_1 MD/SD	SL-ℓ_1 's higher dimension extension, Section IV-B, Algorithm 2, with $\eta = 10^6$, $\mu = 10$
SR-LS	resorts to least square minimization and uses square ranges, [5]
wSR-LS	weighted version of SR-LS, [5]
SDR	a semidefinite relaxation-based method, [9]
LLS	a two-step linearization based algorithm originally developed for localization based on received power (RSSI), [12]
PSDP	SDP based algorithm originally developed for sensor network localization with uncertainty in anchor positions, [13]
ML	the ML estimator initialized with true values for the source localization problem in (1) for $p = 1$; $q = 2$, [6]
ML-lap	the ML estimator (Matlab's function *fminsearch*) initialized with true values for the source localization problem in (1) for $p = 1$; $q = 1$
CRLB	the Cramér-Rao Lower Bound for the source localization problem, Appendix D

We adopt the weighting procedure in [19], [20] that penalizes larger ranges and assumes knowledge of the variances of the disturbances for the weights used in wSR-LS and LLS.

In each simulation we performed M Monte Carlo runs, where in each run the source and anchor locations were randomly generated from a uniform distribution over a square or cube. Observed ranges were generated as described in Section II under appropriate noise densities. The figures show Root Mean-Square Errors (RMSE), computed as $\sqrt{(1/M) \sum_{i=1}^{M} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2}$, where $\mathbf{x}_i$ and $\hat{\mathbf{x}}_i$ denote the actual and estimated source positions in the ith run[2].

Some plots also show the Cramér-Rao Lower Bound (CRLB) for the source location, computed as $\sqrt{(1/M) \sum_{k=1}^{M} \text{tr}(\mathbf{F}_k^{-1})}$, where $\mathbf{F}_k$ denotes the Fisher Information Matrix (FIM) (Appendix D) on the kth Monte Carlo run [11], [21]. Under our nonlinear observation model there is no efficient unbiased estimator, and the bound becomes loose in the presence of large Gaussian perturbations [22]. Still, it is of interest and commonly used as a benchmark. We model disturbances in outlier measurements as Laplacian, selective Gaussian (see below), or Exponentially Modified Gaussian (EMG) [23]–[25] variables. We do not compute the CRLB for Laplacian (non-smooth likelihood) and selective Gaussian (too cumbersome) noise models.

Localization under gaussian noise: We performed $M = 1000$ Monte Carlo runs, where the source and anchor locations were randomly generated from a uniform distribution over a cube whose sides are [0,10]. Observed ranges, corrupted by i.i.d. Gaussian noise with $\sigma_{\text{Gaussian}} \in [10^{-2}, 0.8]$, were generated according to Section II.

Fig. 2(a) shows the RMSE of the algorithms for 6 anchors. The worst performances are attained by SR-LS and LLS. The former squares measurements [$p = 2$, $q = 2$ in (1)], and thus becomes more sensitive to the presence of (Gaussian) noise in range measurements. The latter resorts to linearization, which is not very accurate for this observation model and becomes less so as the noise power increases. The weighted version (wSR-LS) performs significantly better than SR-LS, slightly worse than SLNN, and 7% worse than SL-ℓ_1 MD at high noise levels. wSR-LS achieves these improvements by de-emphasizing long

ranges where the impact of squared disturbances is strongest. However, in practice this method requires some form of calibration to estimate the variances of the disturbances. PSDP[3] and SDR have similar performance and are 7% worse than SL-ℓ_1 MD (28), which is 1.5% better than the simplified version SL-ℓ_1 SD (30).

The results for the various algorithms under Gaussian noise tend to cluster into three performance classes: LLS and SR-LS form a class with relatively large RMSEs, though the latter does considerably better than the former. The second class comprises wSR-LS and SLNN with intermediate and similar performances. Finally, PSDP, SDR, SL-ℓ_1 SD, and SL-ℓ_1 MD form a class with superior performance. Within the second and third classes our algorithms attain the lowest residual errors. The RMSE gap to the CRLB is significant (about 0.3 m for $\sigma_{\text{Gaussian}} = 0.75$), but it seems unlikely that it could be closed by alternative algorithms.

Localization in the presence of outliers: We adopt the same setup for Gaussian noise, except that ranges are now contaminated either by Laplacian noise with $\sigma_{\text{Laplacian}} \in [0.4, 1.8]$, by what we designate as *selective Gaussian noise*, or by EMG noise. Range measurements for the selective Gaussian model are created as $r_i = \|\mathbf{x} - \mathbf{a}_i\| + w_i + |\epsilon|$, where w_i is a Gaussian noise term with $\sigma_{\text{Gaussian}} = 0.1$ that is present in all observations and ϵ is also a Gaussian disturbance, but with higher standard deviation $\sigma_{\text{outlier}} \in [0.4, 1.6]$, that contaminates only one measured range (i.e., $\epsilon = 0$ for all other observations). This statistical model is less tractable than the Laplacian noise model, but we include it in some of our simulations as it more realistically reflects how outliers might occur in real ranging systems. Fig. 2(b) and (c) show RMSEs under the Laplacian and selective Gaussian outlier generation models. There are now more substantial disparities between different classes of algorithms. Broadly, the RMSE curves in these figures can still be divided into 3 classes: As before, LLS and SR-LS attain the highest RMSEs (the latter faring 16% and 25% worse than SL-ℓ_1 MD for Laplacian and selective Gaussian outliers, respectively). The intermediate group is formed by wSR-LS and SLNN which is matched to Gaussian likelihoods (which fares 12% and 20% worse than SL-ℓ_1 MD for Laplacian and selective Gaussian outliers, respectively, and is now outperformed by

---

[2]To avoid cluttering the plots only the performances of selected algorithms are shown, and the average relative performances of the remaining ones are summarized in the text.

[3]The performances of PSDP and SDR are almost identical for all noise levels. Therefore, they are shown by a single curve which is labelled as SDR in all the figures.
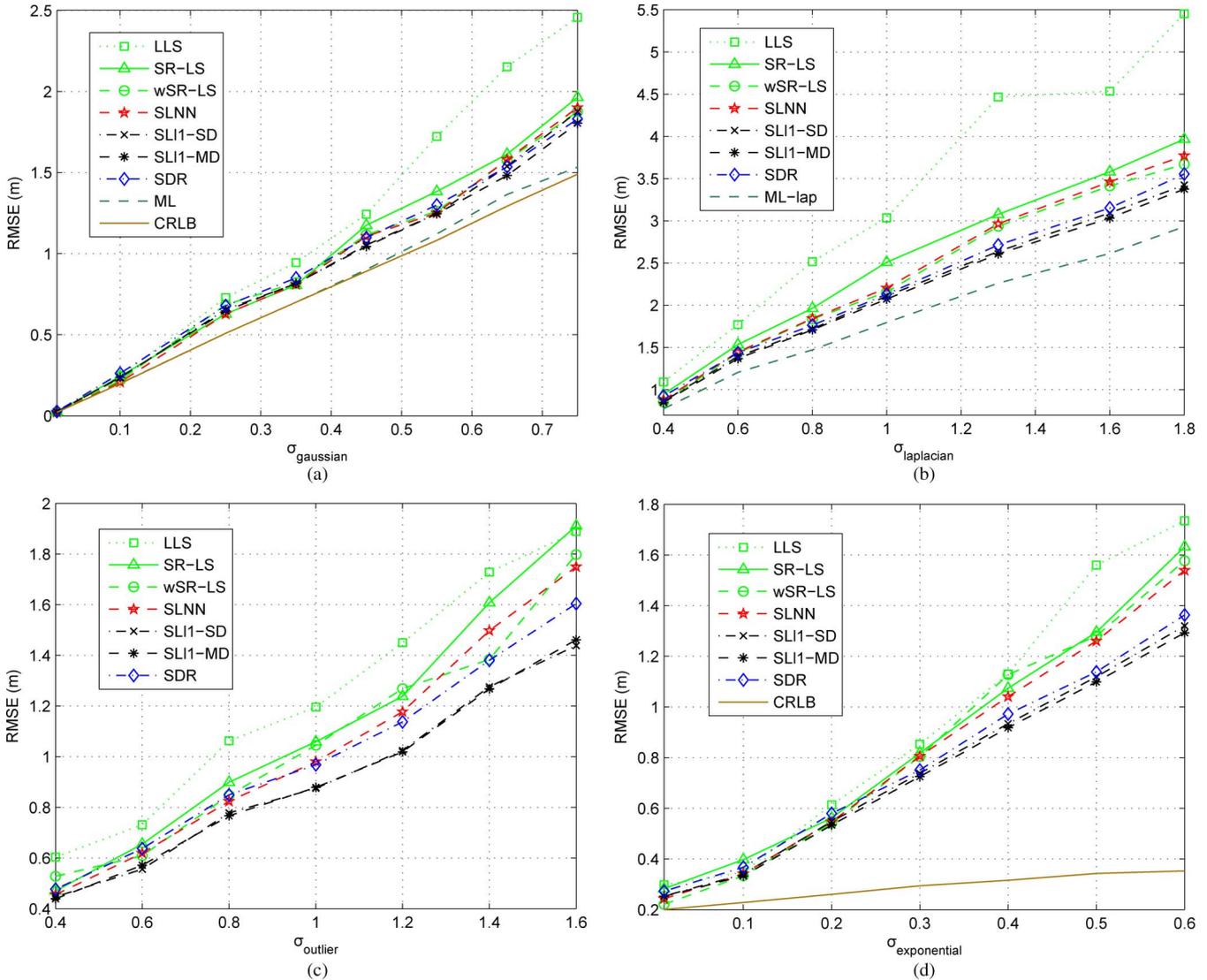
Fig. 2.   3D source localization performance. (a) Gaussian noise; (b) Laplacian outliers; (c) Selective Gaussian outliers; (d) EMG outliers.

PSDP and SDR). Finally, the algorithms designed for Laplacian noise (SL-ℓ_1 MD/SD) outperform all others. Interestingly, their superiority is even more evident under selective Gaussian noise, which does not match the underlying cost (likelihood) function of Section IV-B.

The EMG statistical model is reasonably tractable and often used to model the presence of outliers due to NLOS propagation [23]. Our observations are modeled as i.i.d. and given by $r_i = d_i + b_i + \epsilon_i$, where ϵ_i is a Gaussian noise term with $\sigma_{\mathrm{Gaussian}} = 0.1$ and b_i is an exponential noise term with parameter $\sigma_{\mathrm{exponential}} \in [0.01, 0.6]$. Fig. 2(d) shows that the proposed ℓ_1-based algorithms are still superior under EMG noise. Similarly to the Laplacian and selective Gaussian cases, SR-LS and SLNN fare 22% and 15% worse than SL-ℓ_1 MD, respectively. The gap of all algorithms to the CRLB is now even wider than in the Gaussian case, which could be attributed to looseness of the bound and the fact that the cost functions are not matched to the EMG likelihood function.

Overall, our ℓ_1-based algorithms show the most consistent behavior across all noise models. Also, some enhancements

to LLS used in our simulations are available in the literature, which could considerably reduce the RMSE in the presence of outliers and bring it closer to that of (w)SR-LS [26].

Practical complexity of algorithms vs. ambient space dimension and number of anchors: Several authors have examined the theoretical (worst-case) complexity of localization algorithms, e.g., [11], [27], [28]. The emphasis here is on demonstrating the practical feasibility of our algorithms in centralized scenarios with moderate computational power, so we focus on actual running times, knowing that many technological factors related to hardware and software architectures may influence it[4]. These times are evaluated in 2D and 3D for variable numbers of anchors that encompass those that one would reasonably expect to find within range of a given target position in practice. Our experiments were conducted on a machine

---

[4]Note, however, that SLCP, SLNN, SL $- \ell_1$ MD/SD, as well as PSDP and SDR, are single-shot formulations that essentially require a single invocation of the same general-purpose solver, and the running times for these algorithms should therefore closely reflect the relative operations count.
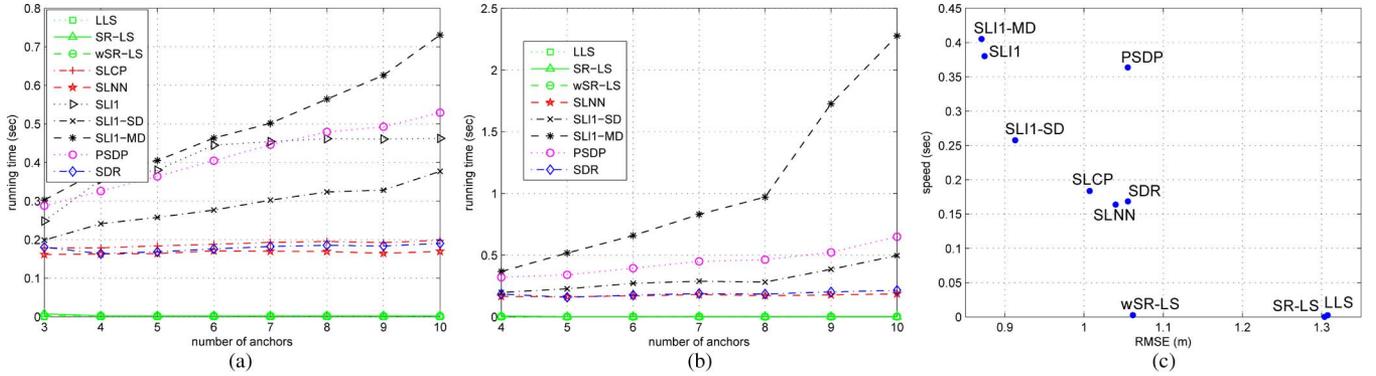
Fig. 3. Speed comparison of algorithms vs. number of anchors. (a) 2D scenario; (b) 3D scenario; (c) Accuracy vs. speed in 2D.

powered by an Intel Xeon 3.2 GHz Quad-Core CPU and 8 GB of RAM, using Matlab R2010b and CVX/SeDuMi as a general-purpose SDP solver. Fig. 3(a) shows that SR-LS, wSR-LS, and LLS are much faster than the other algorithms (average times are 0.2 ms for LLS, 3 ms for SR-LS, and 2.5 ms for wSR-LS), but their larger RMSEs would make them preferred mainly under severe temporal or computational constraints. The running times of SDR, SLCP, and SLNN are almost constant in the figure (about 0.2 s), suggesting that the fixed overhead from our software implementation plays an important role (the variable component becomes significant above 15 anchors). Regarding SL-ℓ_1, SL-ℓ_1 SD and PSDP, the times increase moderately, remain below 0.5 s for the first two algorithms and below 0.7 s for the latter, and are similar in 2D and 3D [Fig. 3(b)]. The largest time and increase rate is attained by SL-ℓ_1 MD. Note that in SL-ℓ_1 MD the number of variables is affected *multiplicatively* by the ambient space dimension, leading to larger effort in 3D.

The above results show that SL-ℓ_1 SD has nearly the same performance of SL-ℓ_1 MD at a fraction of the complexity. The running times for these algorithms, even with more modest hardware, may enable tracking of relatively slow targets with practical numbers of anchors. Fig. 3(c) summarizes the RMSE vs. speed tradeoff of the considered algorithms in 2D for 5 anchors and Selective Gaussian noise with $\sigma_{\text{outlier}} = 1.5$.

Convexity and tightness of SLCP: In this example we characterize the accuracy of the convex relaxation used in SLCP and compare it to the SDR algorithm of [9]. Range measurements to a variable number of randomly placed anchors in a square $[-10, 10] \times [-10, 10]$ were generated over $M = 1000$ Monte Carlo runs, and corrupted by white Gaussian noise.

First, we estimate how often the constraint set $\mathcal{S}(32)$, which appears in our formulation of the source localization problem prior to relaxation (31), is convex along its "upper right" boundary. As discussed in Appendix A, when this property holds the relaxed solution Φ obtained by SLCP (6) will have rank 1 and can be factorized to yield the optimal point for the non-relaxed problem (31) on the boundary of $\mathcal{S}$. We empirically assess convexity of $\mathcal{S}$ by tracing the boundary of the (partially hypothesized) convex hull $\mathcal{T}(33)$ and searching for line segments that delimit regions where the boundaries of $\mathcal{S}$ and $\mathcal{T}$ depart due to local concavity of $\mathcal{S}$. We solve the support hyperplane problem (34) for a grid of angles $0 \leq \beta \leq (\pi/2)$ and detect the presence of a line segment when the distance between the intersection

TABLE II
SOURCE LOCALIZATION ACCURACY FOR RELAXATION-BASED METHODS

σ_{gaussian}	N_{SDR}	N_{SLCP}	SDR [9]	SDRt [9]	SLCP	SLCPt
10^{-3}	490	921	0.0045	0.0014	0.0020	0.0015
10^{-2}	444	815	0.0162	0.0107	0.0112	0.0108
10^{-1}	478	527	0.1503	0.0960	0.1207	0.0959
1	538	526	1.6070	1.1885	1.2169	1.1885

points $(u(\beta), v(\beta))$ for two consecutive angles β exceeds a threshold. For a noise standard deviation $\sigma_{\text{Gaussian}} = 10^{-2}$, $\mathcal{S}$ passed the convexity test in 80% of runs for three anchors. The value increased to 84% for five anchors, in line with our reasoning in Appendix A that $\mathcal{S}$ is more likely to be convex as the number of anchors increases.

Next, we compare the RMSEs of SDR [9] and SLCP. As in [6] we provide results for all Monte Carlo runs (denoted by SDR, SLCP) and also for so-called *tight runs* (denoted by SDRt, SLCPt) where the solution for the relaxed localization problem is close to having rank 1, as desired for subsequent factorization to obtain the actual source coordinates. A solution matrix is considered to be tight when the ratio between its first and second eigenvalues is at least 10^2. Table II lists the RMSEs and the number of tight runs (N_{SDR}, N_{SLCP}) over 1000 trials for five anchors and variable Gaussian noise power. SLCP is clearly superior over the full set of trials, but the gap to SDR closes in the subset of tight runs, indicating that the advantage is mostly due to a much higher probability of its solution having near rank 1. Even for the highest noise power, where the number of tight runs in both algorithms is comparable, the ratio of first to second eigenvalues is usually higher in SLCP, leading to lower RMSE.

VI. CONCLUSION

We have proposed approximate ML algorithms for range-based source localization through convex relaxation of likelihood functions for Gaussian or Laplacian noise. These centralized algorithms exhibit appealing tradeoffs between localization accuracy and computational cost. The fact that they solve a single convex optimization problem and do not require setting critical parameters *a priori* makes them robust and very convenient for practical use, if execution times on the order of 1 second or less using contemporary generic computers can be tolerated.

We have used complex formulations for 2D localization as a starting point, and derived comparable real formulations that can be extended to 3D and higher dimensional scenarios. The flexibility stemming from full 3D localization capability is very relevant in practice, e.g., in indoor localization systems where purely planar anchor/source configurations may be technically cumbersome or altogether unfeasible. As in the 2D formulations (SLCP, SL-ℓ_1), the relaxations underlying our new 3D algorithms (SLNN, SL-ℓ_1 MD, and SL-ℓ_1 SD) are very consistent, yielding matrix solutions that approximately have the rank-1 property with high probability. Our new analysis of the geometry of SLCP sheds light on its observed high accuracy and consistency in various scenarios.

Simulation results show that, for Gaussian noise, our algorithms are more accurate than linearization approaches and those relying on squared ranges. Broadly, their accuracy is comparable to state-of-the-art relaxation-based methods and to (nonrelaxed) ML methods. In the presence of outliers our ℓ_1-based algorithms clearly outperform all the remaining ones, particularly under the selective Gaussian noise model. Overall, SL-ℓ_1 SD emerged as a versatile algorithm that delivers very good performance under different types of disturbance and whose moderate complexity scales favorably with the number of anchors and the ambient space dimension.

APPENDIX A
TIGHTNESS AND GEOMETRY ANALYSIS OF SLCP

One of the salient features of SLCP is the observed tightness of its relaxation, i.e., when the source and anchors are placed randomly the optimal solution is more likely to have nearly rank-1 than other convex relaxations such as [9]. We shed some light on this property by examining the geometry of SLCP's cost function and constraint set, lending support to the technique used in Section V to empirically assess the convexity of specific instances of SLCP by tracing the boundary of the convex hull for the non-relaxed constraint set.

The source localization problem prior to relaxation (5) is

$$
\begin{aligned}
\underset{u,v}{\text{maximize}} \quad & 2\sqrt{u} + \frac{1}{m}v \\
\text{subject to} \quad & (u, v) \in \mathcal{S}, \\
& \mathcal{S} = \left\{ \left(|\mathbf{c}^H \boldsymbol{\theta}|^2, |\mathbf{r}^T \boldsymbol{\theta}|^2 \right) : \boldsymbol{\theta} \in \mathbb{C}^m, |\theta_i| = 1 \right\}.
\end{aligned}
$$
(31)
(32)

The objective function in (31) is concave with respect to u and v, and the optimization problem would be convex if the constraint set $\mathcal{S}$ were convex. Then, the SDR used in SLCP (6) would always find a desired rank-1 solution $\boldsymbol{\Phi}$.

Given the separable form of the cost function (31) it is clear that, for fixed v, it can be maximized by choosing u as large as possible within $\mathcal{S}$, and vice-versa. This implies the following:

Property 1: Optimal points of (31) lie on the "upper right" boundary of set $\mathcal{S}$, i.e., they are maximal elements of $\mathcal{S}$ with respect to the standard cone $\mathbb{R}_+^2$ [16].

Regarding the convexity properties of $\mathcal{S}$, first recall that the cost function of (31) was designed to be invariant to rotations of $\boldsymbol{\theta}$, so its first element may be taken as unity. For $m = 2$ anchors and $\theta_1 = 1$, $\theta_2 = e^{j\phi}$ set $\mathcal{S}$ as defined in (32) is readily seen to be an ellipse. Given the definitions of $\mathbf{c}$ and $\mathbf{r}$ in (4), for
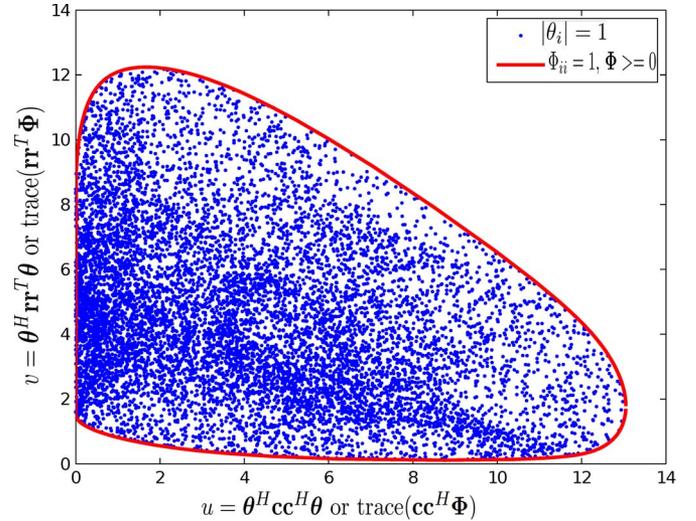

Fig. 4. Constraint set $\mathcal{S}$ for 4 anchors and randomly generated $\boldsymbol{\theta}$ satisfying $|\theta_i| = 1$. The hypothesized convex hull, computed by relaxation of $\mathcal{S}$, is also depicted.

$m > 2$ anchors it is always possible to effectively zero out the contributions from $\theta_3, \ldots, \theta_m$ when $r_3 = \ldots = r_m = 0$ in the diagonal of $\mathbf{R}$, thus reverting to the case $m = 2$. Hence, set $\mathcal{S}$ may be nonconvex for any number of anchors.

In spite of the lack of convexity guarantees for $\mathcal{S}$, our simulation results suggest that for $m \geq 3$ anchors and typical range measurements this set usually does have a convex-like shape. Even when $\mathcal{S}$ is not convex all that is required for our SDR to provide a rank-1 solution is "local convexity" along the "upper right" boundary of $\mathcal{S}$ where the optimal point of (31) is known to be located. More formally, we require that the intersection of $\mathcal{S}$ with any supporting hyperplane defined by a normal direction with nonnegative components be a compact subset (a single point or a line segment) [16]. Fig. 4 depicts an example of $\mathcal{S}$ for $m = 4$ anchors and randomly generated $\mathbf{c}, \mathbf{r}$. Our practical test for local (non)convexity of $\mathcal{S}$ consists of tracing multiple supporting hyperplanes with nonnegative normal elements, and assessing whether any of them intersect $\mathcal{S}$ at two well-separated points. We build supporting hyperplanes not on $\mathcal{S}$ directly, which is a hard problem, but on the related relaxed convex set

$$
\mathcal{T} = \left\{ \left(\text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\Phi}), \text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\Phi}) \right) : \boldsymbol{\Phi} \in \mathbb{C}^{m \times m}, \boldsymbol{\Phi} \succeq 0, \phi_{ii} = 1 \right\}.
$$
(33)

Specifically, for a supporting hyperplane with normal $(\cos \beta, \sin \beta)$, $0 \leq \beta \leq (\pi/2)$, we determine an intersection point with $\mathcal{T}$ by solving the convex optimization problem

$$
\begin{aligned}
\underset{\boldsymbol{\Phi}}{\text{maximize}} \quad & \left\langle (\cos \beta, \sin \beta), \left(\text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\Phi}), \text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\Phi}) \right) \right\rangle \\
\text{subject to} \quad & \boldsymbol{\Phi} \succeq 0, \phi_{ii} = 1,
\end{aligned}
$$
(34)

and setting the intersection as $u = \text{tr}(\mathbf{c}\mathbf{c}^H \boldsymbol{\Phi})$, $v = \text{tr}(\mathbf{r}\mathbf{r}^T \boldsymbol{\Phi})$. This is justified by the following lemma (see Appendix B).

Lemma 1: For $m \leq 3$ anchors the sets $\mathcal{S}$ and $\mathcal{T}$ have the same set of supporting hyperplanes with nonnegative normal elements. Equivalently, in the relevant portion of its boundary $\mathcal{T}$ coincides with the convex hull of $\mathcal{S}$.

Although we only prove this result up to $m = 3$, the empirical evidence suggests that it is also valid for higher m and arbitrary supporting hyperplanes (i.e., $\mathcal{T}$ is actually the convex hull of $\mathcal{S}$). We leave this as a conjecture and apply the procedure for $m > 3$ as well, noting, however, that the case $m = 3$ has major practical significance as the minimum number of anchors that are necessary to recover a general 2D source position based on range measurements.

APPENDIX B
PROOF OF LEMMA 1

Our proof of Lemma 1 relies on a result, interesting in its own right, that characterizes the convex hull of the set of 3×3 rank-1 matrices built from complex vectors with unit-magnitude components.

Lemma 2: Let

$$\mathcal{A} = \left\{ \boldsymbol{\theta}\boldsymbol{\theta}^H : \boldsymbol{\theta} \in \mathbb{C}^3, |\theta_i| = 1 \right\}, \tag{35}$$

$$\mathcal{B} = \left\{ \boldsymbol{\Phi} \in \mathbb{C}^{3\times3} : \boldsymbol{\Phi} \succeq 0, \phi_{ii} = 1 \right\} \tag{36}$$

then $\mathcal{B} = \text{co}(\mathcal{A})$.

Proof: $\text{co}(\mathcal{A}) \subset \mathcal{B}$ is straightforward since $\mathcal{B}$ is convex and $\mathcal{A} \subset \mathcal{B}$. For the reverse direction $\text{co}(A) \supset \mathcal{B}$ our goal is to find, for every $\boldsymbol{\Phi} \in \mathcal{B}$, matrices $\boldsymbol{\Theta}_i \in \mathcal{A}$ and nonzero scalars $\lambda_i \geq 0$, with $\sum_i \lambda_i = 1$, such that $\boldsymbol{\Phi} = \sum_i \lambda_i \boldsymbol{\Theta}_i$.

Note that both $\mathcal{A}$ and $\mathcal{B}$ are invariant under the (unitary) similarity operation, $\mathbf{M} \to \mathbf{PMP}^H$, where $\mathbf{P}$ is the product of a permutation and a diagonal unitary matrix. In other words, we can simultaneously permute rows and columns and multiply the ith row and ith column by a unit-magnitude complex number. Thus, we can assume without loss of generality that $\boldsymbol{\Phi}$ is of the form

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & a & b \\ a & 1 & z^* \\ b & z & 1 \end{bmatrix}, \quad 0 \leq a \leq b, \quad z \in \mathbb{C}. \tag{37}$$

Since $\boldsymbol{\Phi} \succeq 0$, we must have $a \leq 1$, $b \leq 1$, $|z| \leq 1$, and $0 \leq |\boldsymbol{\Phi}| = 1 - a^2 - b^2 - |z|^2 + 2ab \, \text{Re}\{z\}$, which, for $z = x + jy$, reads

$$(x - ab)^2 + y^2 \leq (1 - a^2)(1 - b^2). \tag{38}$$

For fixed a, b this inequality describes a circle (with interior) in the (x, y) plane, centered on $(ab, 0)$. Since any point in the interior of a circle can be written as a convex combination of two points on its boundary, we can assume that we have equality in (38). Thus, from now on we assume

$$z = ab + \sqrt{(1 - a^2)(1 - b^2)}e^{j\varphi}. \tag{39}$$

We now complete the proof by expressing such $\boldsymbol{\Phi}$ as a convex combination of two matrices from $\mathcal{A}$. For given $0 \leq a \leq b \leq 1$ and $\varphi \in [0, 2\pi[$ we want to find $\alpha, \beta, \gamma, \delta \in [0, 2\pi[$, and $0 \leq \lambda \leq 1$ such that (37) equals

$$\lambda \begin{bmatrix} 1 \\ e^{j\alpha} \\ e^{j\beta} \end{bmatrix} [1 \; e^{-j\alpha} \; e^{-j\beta}] + (1 - \lambda) \begin{bmatrix} 1 \\ e^{j\gamma} \\ e^{j\delta} \end{bmatrix} [1 \; e^{-j\gamma} \; e^{-j\delta}]. \tag{40}$$
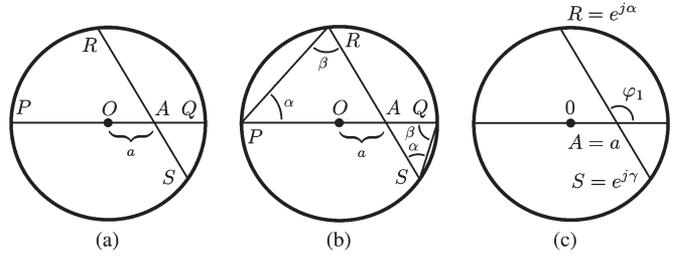

Fig. 5. Illustration of geometrical Lemma 3 and its application to a convex combination on the unit circle.

We thus have

$$a = \lambda e^{j\alpha} + (1 - \lambda)e^{j\gamma}, \quad b = \lambda e^{j\beta} + (1 - \lambda)e^{j\delta}, \tag{41}$$

$$z = \lambda e^{j(\beta-\alpha)} + (1 - \lambda)e^{j(\delta-\gamma)}. \tag{42}$$

From the first two relations we get $e^{j\gamma} = (a - \lambda e^{j\alpha})/(1 - \lambda)$, $e^{j\delta} = (b - \lambda e^{j\beta})/(1 - \lambda)$, and replacing these in (42) yields, after simple manipulations

$$z = ab + \frac{\lambda}{1 - \lambda}(e^{-j\alpha} - a)(e^{j\beta} - b). \tag{43}$$

Before proceeding, we state and prove a useful lemma from elementary geometry:

Lemma 3: Referring to Fig. 5(a), if A is a point inside a unit circle whose distance to the center is a, RS is any line through A, and PQ is a diameter through A, then

$$AR \cdot AS = AP \cdot AQ = (1 - a)(1 + a) = 1 - a^2. \tag{44}$$

This is proved by noting that triangles APR and AQS, depicted in Fig. 5(b), are similar, hence $AP/AS = AR/AQ$. We use the lemma above with parameters as depicted in Fig. 5(c). From $A = \lambda R + (1 - \lambda)S$ we have $(AR/AS) = (1 - \lambda)/\lambda$, and by Lemma 3 $AR \cdot AS = 1 - a^2$, hence

$$AR = \sqrt{\frac{1 - \lambda}{\lambda}(1 - a^2)}, \quad e^{j\alpha} = a + \sqrt{\frac{1 - \lambda}{\lambda}(1 - a^2)}e^{j\varphi_1}. \tag{45}$$

Similarly, with $A = b$, $R = e^{j\beta}$, $S = e^{j\delta}$, and φ_2 instead of φ_1, we have

$$e^{j\beta} = b + \sqrt{\frac{1 - \lambda}{\lambda}(1 - b^2)}e^{j\varphi_2}. \tag{46}$$

Substituting (45), (46) back in (43) yields $z = ab + \sqrt{(1 - a^2)(1 - b^2)}e^{j(\varphi_2-\varphi_1)}$, which has the same form as (39), obtained from the positive semidefinite condition for matrix $\boldsymbol{\Phi}$ in (37).

We now argue that letting angle α go from 0 to 2π is equivalent to letting φ_1 cover an interval of length 2π as well [Fig. 5(c)]. Fixing φ_1, and consequently α, the two relations in (41), together with an arbitrary requirement that $\text{Im}\{e^{j\beta}\} \geq 0$, fix the values[5] of β, γ, δ, λ, and, in particular, of φ_2. Thus, $\varphi_2 = f(\varphi_1)$ is a continuous function of φ_1.

---

[5] Equivalently, fixing φ_1 defines the geometrical construction shown in Fig. 5(c), setting γ and λ. Then, λ defines the second construction for $A = b$ if $\text{Im}\{e^{j\beta}\} \geq 0$ is additionally specified, setting β, δ, and φ_2.

When $\varphi_1 = 0$, φ_2 has a certain value, say, $\varepsilon_0 \in [0, \pi]$ (it can be computed, but is not needed in this proof). For $\varphi_1 = \pi$ it is straightforward to see that $\varphi_2 = \pi - \varepsilon_0$, and for $\varphi_1 = 2\pi$ it is again ε_0. In particular, the continuous function $\varphi_2 - \varphi_1$ takes values from $\varepsilon_0 - 0 = \varepsilon_0$ to $\varepsilon_0 - 2\pi$, i.e., modulo 2π it takes all values in $[0, 2\pi[$. Thus, for any given angle φ in (39), let φ_1 be such that $f(\varphi_1) - \varphi_1 = \varphi$, modulo 2π. Then, the corresponding α, β, γ, δ, and λ, as explained above, give the desired decomposition (40). ∎

We proceed and prove Lemma 1 under the assumption of Lemma 2, thus assuming $m = 3$.

Proof: We rewrite sets $\mathcal{S}$ in (32) and $\mathcal{T}$ in (33) using the notation (35)

$$\mathcal{S} = \left\{ (\mathbf{c}^H \boldsymbol{\Theta} \mathbf{c}, \mathbf{r}^T \boldsymbol{\Theta} \mathbf{r}) : \boldsymbol{\Theta} \in \mathcal{A} \right\}, \tag{47}$$

$$\mathcal{T} = \left\{ (\mathbf{c}^H \boldsymbol{\Phi} \mathbf{c}, \mathbf{r}^T \boldsymbol{\Phi} \mathbf{r}) : \boldsymbol{\Phi} \in \mathrm{co}(\mathcal{A}) \right\}. \tag{48}$$

Obviously $\mathcal{S} \subset \mathcal{T}$. Now let $\alpha \in [0, (\pi/2)]$ and define

$$(u_1, v_1) = \arg \max_{(u,v) \in \mathcal{T}} \langle (\cos \alpha, \sin \alpha), (u, v) \rangle. \tag{49}$$

We wish to show that

$$\langle (\cos \alpha, \sin \alpha), (u_1, v_1) \rangle = \max_{(u,v) \in \mathcal{S}} \langle (\cos \alpha, \sin \alpha), (u, v) \rangle \tag{50}$$

so that the inner product over $\mathcal{S}$ attains the same maximum as over the larger set $\mathcal{T}$, and the support hyperplanes with normal $(\cos \alpha, \sin \alpha)$ coincide for the two sets. It is enough to prove that there exists $(u', v') \in \mathcal{S}$ that attains the left-hand side of (50). We may write $\boldsymbol{\Phi}_1 \in \mathrm{co}(\mathcal{A})$ which maximizes (49) as $\boldsymbol{\Phi}_1 = \sum_i \lambda_i \boldsymbol{\theta}_i \boldsymbol{\theta}_i^H, \lambda_i \geq 0, \sum_i \lambda_i = 1, |\theta_{ik}| = 1$, hence the left-hand side of (50) is given by

$$(\sqrt{\cos \alpha}\, \mathbf{c})^H \boldsymbol{\Phi}_1 (\underbrace{\sqrt{\cos \alpha}\, \mathbf{c}}_{\mathbf{p}}) + (\sqrt{\sin \alpha}\, \mathbf{r})^T \boldsymbol{\Phi}_1 (\underbrace{\sqrt{\sin \alpha}\, \mathbf{r}}_{\mathbf{q}})$$
$$= \sum_i \lambda_i \left(|\mathbf{p}^H \boldsymbol{\theta}_i|^2 + |\mathbf{q}^H \boldsymbol{\theta}_i|^2 \right). \tag{51}$$

Let i_0 be the index for the maximum term in the sum. Then $\langle (\cos \alpha, \sin \alpha), (u_1, v_1) \rangle \leq |\mathbf{p}^H \boldsymbol{\theta}_{i_0}|^2 + |\mathbf{q}^H \boldsymbol{\theta}_{i_0}|^2 = \langle (\cos \alpha, \sin \alpha), (\mathbf{c}^H \boldsymbol{\theta}_{i_0} \boldsymbol{\theta}_{i_0}^H \mathbf{c}, \mathbf{r}^T \boldsymbol{\theta}_{i_0} \boldsymbol{\theta}_{i_0}^H \mathbf{r}) \rangle$, which completes the proof because the second argument in the inner product is an element of $\mathcal{S}$. ∎

APPENDIX C
ANALYSIS OF SLNN

Solution of the inner subproblem (12): We reformulate (12) as

$$\begin{aligned} \underset{\mathbf{V}}{\text{minimize}} \quad & \frac{1}{2}\|\mathbf{V} - \mathbf{B}\|_F^2 \\ \text{subject to} \quad & \mathbf{V}^T \mathbf{V} = \mathbf{I}_n, \end{aligned} \tag{52}$$

with $\mathbf{B} = -\mathbf{U}^T \mathbf{C}$, whose cost function differs from (12) by a constant on the constraint set of orthogonal matrices and therefore has the same minimizer. Recognizing (52) as an orthogonal Procrustes problem [29], its solution for this particular instance

is known to be obtained from the SVD of $\mathbf{B} = \mathbf{S} \boldsymbol{\Sigma} \mathbf{W}^T$ as $\mathbf{V} = \mathbf{S} \mathbf{W}^T$. Inserting this into (12) yields the optimal cost

$$\mathrm{tr}(-\mathbf{B}^T \mathbf{V}) = -\mathrm{tr}\left(\mathbf{W} \boldsymbol{\Sigma}^T \underbrace{\mathbf{S}^T \mathbf{S}}_{\mathbf{I}} \mathbf{W}^T \right) = -\mathrm{tr}\left(\boldsymbol{\Sigma}^T \underbrace{\mathbf{W}^T \mathbf{W}}_{\mathbf{I}} \right)$$
$$= -\mathrm{tr}(\boldsymbol{\Sigma}^T) = -\|\mathbf{C}^T \mathbf{U}\|_N.$$

∎

Proof of equivalence between (15) and (16): We first rewrite (16) using an equivalent Schur complement

$$\begin{aligned} \underset{\mathbf{W}, \mathbf{Z}}{\text{maximize}} \quad & 2\,\mathrm{tr}(\mathbf{Z}) + \frac{1}{m}\mathrm{tr}(\mathbf{r}\mathbf{r}^T \mathbf{W}) \\ \text{subject to} \quad & \mathbf{W} \succeq 0, \quad w_{ii} = 1, \quad \mathbf{Z}^2 \preceq \mathbf{C}^T \mathbf{W} \mathbf{C}, \quad \mathbf{Z} \succeq 0. \end{aligned} \tag{53}$$

Let p_1^* and p_2^* be the optimal values of problems (15) and (53), respectively. Choose a feasible point $(\mathbf{Z}, \mathbf{W})$ for the second problem, such that $0 \preceq \mathbf{Z}^2 \preceq \mathbf{C}^T \mathbf{W} \mathbf{C}$. This implies[6] $\mathbf{Z} \preceq (\mathbf{C}^T \mathbf{W} \mathbf{C})^{1/2}$, hence the values of the objective functions satisfy

$$2\,\mathrm{tr}(\mathbf{Z}) + \frac{1}{m}\mathrm{tr}(\mathbf{r}\mathbf{r}^T \mathbf{W}) \leq 2\,\mathrm{tr}\left((\mathbf{C}^T \mathbf{W} \mathbf{C})^{\frac{1}{2}} \right) + \frac{1}{m}\mathrm{tr}(\mathbf{r}\mathbf{r}^T \mathbf{W}). \tag{54}$$

Choosing for $(\mathbf{Z}, \mathbf{W})$ the unique maximizer of (53), inequality (54) asserts that $p_1^* \geq p_2^*$. For the converse choose a feasible point $\mathbf{W}$ for the first problem and consider the eigendecomposition $\mathbf{C}^T \mathbf{W} \mathbf{C} = \mathbf{Q} \boldsymbol{\lambda} \mathbf{Q}^T$. Now set $\mathbf{Z} = \mathbf{Q} \boldsymbol{\lambda}^{1/2} \mathbf{Q}^T$, so that $\mathbf{Z}^2 = \mathbf{Q} \boldsymbol{\lambda} \mathbf{Q}^T = \mathbf{C}^T \mathbf{W} \mathbf{C}$, and $(\mathbf{W}, \mathbf{Z})$ is therefore feasible for (53). For both problems the value of the cost function is $2\,\mathrm{tr}(\boldsymbol{\lambda}^{1/2}) + (1/m)\mathrm{tr}(\mathbf{r}\mathbf{r}^T \mathbf{W})$. Choosing for $\mathbf{W}$ the maximizer of (15) the construction for $\mathbf{Z}$ yields a feasible point $(\mathbf{W}, \mathbf{Z})$ for (53) with cost p_1^*. Therefore, $p_1^* \leq p_2^*$, and coupling this with the converse inequality above we conclude that $p_1^* = p_2^*$ and the two problems are equivalent. ∎

APPENDIX D
CRLB FOR SOURCE LOCALIZATION BASED ON TOA

Several authors have addressed the problem of deriving performance bounds for source localization based on different types of observations under both LOS and NLOS propagation [23]–[25]. As pointed out in [24] the FIM for the localization problem can be expressed, under mild conditions, as the product of a scalar defined by the noise statistics alone, and a matrix that depends only on the geometry of the source and anchors

$$\mathbf{F} = g \sum_{i=1}^{m} \mathbf{F}_i \tag{55}$$

where

$$\mathbf{F}_i = \frac{(\mathbf{x} - \mathbf{a}_i)(\mathbf{x} - \mathbf{a}_i)^T}{\|\mathbf{x} - \mathbf{a}_i\|^2}, \tag{56}$$

$$g = \mathbb{E}\left\{ \left(\frac{\partial}{\partial w} \ln f(w) \right)^2 \right\} = \int_{-\infty}^{\infty} \frac{f'(w)^2}{f(w)}\, dw, \tag{57}$$

---

[6] $\mathbf{A} \succeq \mathbf{B} \succeq 0 \Rightarrow \mathbf{A}^{1/2} \succeq \mathbf{B}^{1/2} \succeq 0$ [29].

and f denotes the additive noise pdf that affects range measurements as described at the start of Section II. The assumptions underlying (55) are that f is differentiable over its support region, say, $[a, b]$, and that $f(a) - f(b) = 0$. These are satisfied by many distributions, including EMG [25], [30], Rayleigh [25], and several types of mixtures that are often used to model NLOS propagation (but not by the Laplacian pdf that underlies our ℓ_1-based algorithms). In the LOS case f is Gaussian $\mathcal{N}(0, \sigma^2)$ and the constant above evaluates to $g = \sigma^{-2}$, such that (55) coincides with the well-known Gaussian bound.

In the EMG model measurements are corrupted by the sum of independent Gaussian (mean 0, variance σ^2) and exponential (reciprocal rate μ) components. The density is the convolution of both pdfs, given by $f(w) = (1/2\mu) \exp(-(w/\mu) + (\sigma^2/2\mu^2)) \text{erfc}(-(w/\sqrt{2}\sigma) + (\sigma/\sqrt{2}\mu))$, with mean value μ and variance $\sigma^2 + \mu^2$. We evaluate the integral (57) numerically with sufficiently large upper and lower limits.

REFERENCES

[1] W. Dargie and C. Poellabauer, *Fundamentals of Wireless Sensor Networks: Theory and Practice*. Hoboken, NJ, USA: Wiley, 2010.

[2] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proc. ACM MobiCom*, San Diego, CA, USA, Sep. 2003, pp. 81–95.

[3] G. Wang and K. Yang, "A new approach to sensor node localization using RSS measurements in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, pp. 1389–1395, May 2011.

[4] P. Oğuz-Ekim, J. Gomes, J. Xavier, and P. Oliveira, "A convex relaxation for approximate maximum-likelihood 2D source localization from range measurements," in *Proc. IEEE ICASSP*, Dallas, TX, USA, Mar. 2010, pp. 2698–2701.

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

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

[7] E. Xu, Z. Ding, and S. Dasgupta, "Wireless source localization based on time of arrival measurement," in *Proc. IEEE ICASSP*, Dallas, TX, USA, Mar. 2010, pp. 2842–2845.

[8] Y. Chan, H. Hang, and P. Ching, "Exact and approximate maximum likelihood localization algorithms," *IEEE Trans. Signal Process.*, vol. 55, no. 1, pp. 10–16, Jan. 2006.

[9] 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. IEEE ICASSP*, Montreal, QC, Canada, May 2004, pp. II-145–II-148.

[10] P. Oğuz-Ekim, J. Gomes, J. Xavier, and P. Oliveira, "Robust localization of nodes and time-recursive tracking in sensor networks using noisy range measurements," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3930–3942, Aug. 2011.

[11] E. Xu, Z. Ding, and S. Dasgupta, "Source localization in wireless sensor networks from signal time of arrival measurements," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2887–2897, Jun. 2011.

[12] H. C. So and L. Lin, "Linear least squares approach for accurate received signal strength based source localization," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 4035–4040, Aug. 2011.

[13] W.-K. Lui, W.-K. Ma, H. C. So, and F. K. W. Chan, "Semi-definite programming algorithms for sensor network node localization with uncertainties in anchor positions and/or propagation speed," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 752–763, Feb. 2009.

[14] S. Zhu and Z. Ding, "A simple approach of range-based positioning with low computational complexity," *IEEE Trans. Wireless Commun.*, vol. 8, no. 12, pp. 5832–5836, Dec. 2009.

[15] A. O. Hero and D. Blatt, "Sensor network source localization via Projection Onto Convex Sets (POCS)," in *Proc. IEEE ICASSP*, Philadelphia, PA, USA, Mar. 2005, vol. 3, pp. 689–692.

[16] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[17] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization: Analysis, Algorithms, Engineering Applications*. Philadelphia, PA, USA: SIAM, 2001.

[18] P. Biswas and Y. Ye, "Semidefinite programming for *ad hoc* wireless sensor network localization," in *Proc. IPSN*, Berkeley, CA, USA, Apr. 2004, pp. 46–54.

[19] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Process.*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.

[20] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, "Least squares algorithms for time of arrival based mobile location," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1121–1128, Apr. 2004.

[21] N. Patwari, A. Hero, III, M. Perkins, N. Correal, and R. O. Dea, "Relative location estimation in wireless sensor networks," *IEEE Trans. Signal Process.*, vol. 51, no. 8, pp. 2137–2148, Aug. 2003.

[22] A. Host-Madsen, "On the existence of efficient estimators," *IEEE Trans. Signal Process.*, vol. 48, no. 11, pp. 3028–3031, Nov. 2000.

[23] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position error bound for UWB localization in dense cluttered environments," in *Proc. IEEE ICC*, Istanbul, Turkey, Jun. 2006, pp. 613–628.

[24] V. N. Ekambaram, K. Ramchandran, and R. Sengupta, "Scaling laws for cooperative node localization in non-line-of-sight wireless networks," in *Proc. IEEE GLOBECOM*, Houston, TX, USA, Dec. 2011, pp. 1–5.

[25] J. Huang, P. Wang, and Q. Wan, "CRLBs for WSNs localization in NLOS environments," *EURASIP J. Wireless Commun. Netw.*, vol. 2011, no. 1, p. 16, Jun. 2011.

[26] I. Guvenc and C.-C. Chong, "A survey on TOA based wireless localization and NLOS mitigation techniques," *IEEE Commun. Surveys Tuts.*, vol. 11, no. 3, pp. 107–124, 2009.

[27] G. Destino and G. Abreu, "On the maximum likelihood approach for source and network localization," *IEEE Trans. Signal Process.*, vol. 59, no. 10, pp. 4954–4970, Oct. 2011.

[28] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, Mar. 1996.

[29] R. A. Horn and C. A. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.

[30] K. T. Feng, C. L. Chen, and C. H. Q. Chen, "GALE: An enhanced geometry assisted location estimation algorithm for NLOS environments," *IEEE Trans. Mobile Comput.*, vol. 7, no. 2, pp. 199–213, Feb. 2008.

Pinar Oğuz-Ekim (S'10) received the B.S. and M.S. degrees in electrical and electronics engineering from Middle East Technical University, Ankara, Turkey, in 2003 and 2006, respectively. She received the Ph.D. degree in electrical and computer engineering from Instituto Superior Técnico, Lisbon, Portugal, in 2013. Her Ph.D. thesis won the 2012 IBM Science Prize in Portugal. She is currently an Instructor at the Gebze Institute of High Technology, Kocaeli, Turkey. Her research interests include localization and target tracking in wireless sensor networks and optimization theory.

João Pedro Gomes (S'95–M'03) received the Diploma, M.S. and Ph.D. degrees in electrical and computer engineering from Instituto Superior Técnico (IST), Lisbon, Portugal, in 1993, 1996 and 2002, respectively. He joined the Department of Electrical and Computer Engineering, IST in 1995, where he is currently an Assistant Professor. Since 1994, he has also been a Researcher in the Signal and Image Processing Group, Institute for Systems and Robotics, Lisbon. His current research interests include channel identification and equalization in wireless communications, underwater communications and acoustics, localization algorithms, and sensor networks. Dr. Gomes currently serves as an Associate Editor for signal processing and communications in the *IEEE Journal of Oceanic Engineering*.

João Xavier (S'97–M'03) received the Ph.D. degree in electrical and computer engineering from Instituto Superior Técnico (IST), Lisbon, Portugal, in 2002. Currently, he is an Assistant Professor in the Department of Electrical and Computer Engineering, IST. He is also a Researcher at the Institute for Systems and Robotics, Lisbon, Portugal. His current research interests are in the area of optimization, sensor networks, and signal processing on manifolds.

Paulo Oliveira (S'91–M'95) received the Ph.D. degree from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 2002. He is an Associate Professor in the Department of Mechanical Engineering, IST, Universidade de Lisboa, Lisbon, Portugal, and Senior Researcher in the Institute for Systems and Robotics of LARSyS. His areas of interest are in robotics and autonomous vehicles with special focus on the fields of sensor fusion, navigation, positioning, and signal processing. He has participated in more than 15 Portuguese and European research projects, over the last 20 years and published approximately 200 papers in top journals and conferences.

Marko Stošić received the Ph.D. degree in mathematics from Instituto Superior Tecnico, Lisbon, Portugal, in 2006. Since 2006 he has been a Researcher at the Instituto de Sistemas e Robotica (ISR), Lisbon, and has been working on mathematical problems motivated by signal processing and computer vision. In 2008, he won the Best Innovative Paper Award at the 2008 ICIP Conference. Apart from the applied mathematics, he has a wide range of interests in mathematics and has numerous publications in various different fields of mathematics and physics, including: geometry and topology—most notably in knot theory and categorifications, mathematical physics, matrix theory and control theory.