Algebraic Approach for Robust Localization with Heterogeneous Information

Davide Macagnano, *Member, IEEE*,
and Giuseppe Thadeu Freitas de Abreu, *Senior Member, IEEE*

***Abstract*—We offer a redesigned form of the classical multidimensional scaling (C-MDS) algorithm suitable to handle the localization of multiple sources under line-of-sight (LOS) and non-line-of-sight (NLOS) conditions. To do so we propose to modify the kernel matrix used in the MDS algorithm to allow for both distance and angle information to be processed algebraically (without iteration) and simultaneously. In so doing we also show that the new formulation overcomes two well known limitations of the C-MDS approach, namely the propagation error problem and the possibility to weight the dissimilarities used as measurement information, including, for the case of binary weights, the data erasure problem. Due to the increased size of the proposed edge kernel matrix $\mathbf{K_E}$ used in the algorithm, the Nyström approximation is applied to reduce the overall computational complexity to few matrix multiplications. Range only scenarios are also dealt with by approximating the matrix $\mathbf{K_E}$. Simulations in range-angle as well as range-only scenarios demonstrate the superiority of our solution under both LOS and NLOS conditions versus semidefinite programming (SDP) formulations of the problem specifically designed to exploit the heterogeneity of the information available.**

***Index Terms*—Localization, multidimensional scaling (MDS), heterogeneous networks, range-based measurements, angle-based measurements.**

I. Introduction

In the last decade the need for highly personalized information services has thriven the creation of novel location based services (LBS), contributing to make accurate location information anywhere anytime to become a strategic asset for companies operating in the information technology (IT) business.

All this interest from companies stimulated a surge of interest from the research community. Considerable progress has been made on various aspects of the problem including advanced multilateration techniques [1], synchronization problems [2], increasing accuracy by cooperative schemes [3] and robust ranging in the presence of non-line-of-sight (NLOS) conditions and interference, sophisticated optimization techniques capable to exploit geometric and statistical information characterizing the scenario of interest [4], [5] and distributed solution [6], [7]. In addition to contributions of such nature, it has also been shown that due to its nonparametric formulation, the multidimensional scaling (MDS) can be an advantageous technique compared to more conventional methods especially when the targets mobility cannot be simply and reliably modeled a priori [8], [9]. MDS-based solution applied to the localization problem typically rely on range dissimilarities that can be estimated either on the bases of time of arrival (TOA) or received signal strength (RSS) measurements. Due to the fact that RSS-based ranging is highly subject to environmental conditions such as fading, shadowing and antenna distortion, time-based ranging methods are particularly attractive, especially since the advent of Ultra-Wideband Impulse Radio (UWB-IR) [10].

Another type of measurement used in localization systems is the angle information obtained by estimating the angle of arrival (AOA) of signals at different receivers.

While AOA estimates can be accurately and inexpensively performed with a low-frequecy narrowband signal even under NLOS conditions, provided that the multipath in the environment is not too dense [11], TOA-based ranging can be performed at low cost even in dense multipath environments, provided that line-of-sight (LOS) conditions exist. As a consequence, AOA and TOA measurements can be seen as complimentary information in which the errors affecting the observations are due to entirely different factors. Moreover, as shown in [4, pp. 217] recent systems are able to make both, angle and range observations simultaneously available, which in turn calls for algorithms that can exploit this heterogeneity in the information to enhance the localization accuracy.

The challenge is therefore to design a localization algorithms that can jointly and efficiently exploit angle and ranging information. This paper proposes one such algorithm where angle and ranging information are processed jointly in a natural and straightforward manner. Our contribution is a MDS-based formulation of the localization problem, where the entities are the *vectors* pointing from a source to another, and the measure of dissimilarity their inner product.

In a Graph-Theoretical setting, the technique proposed in [12] translates to modeling the network as an undirected graph $G_{\eta,N}(\mathrm{X}, \vec{\mathrm{V}}, \mathrm{D})$, embedded in a η-dimensional space, where the entities are its N vertices V, and the dissimilarities its weights D. Conversely, in our formulation, the network is modeled as a complete oriented graph (also know as a *tournament*), where the entities are the *edges* V of the graph, and the dissimilarity metric their inner product.[1]

In allusion to its construction, the Gram Kernel used in the

Manuscript received February 17, 2013; revised May 30, 2013; accepted July 15, 2013. The associate editor coordinating the review of this paper and approving it for publication was W. Choi.

Parts of this work were presented at the *7-th IEEE WCNC* in 2007 and at the *45-th IEEE ASILOMAR Conference on Signals, Systems and Computers* in November 2011.

D. Macagnano is with the Centre for Wireless Communications, Department of Communications Engineering, University of Oulu, P.O.Box 4500, 90014-FIN (e-mail: macagnan@ee.oulu.fi).

G. Abreu is with the School of Eng. & Science - SES, Department Elec. Eng. and Computer Science, Jacobs University Bremen, 28759-Germany (e-mail: g.abreu@jacobs-university.de).

Digital Object Identifier 10.1109/TWC.2013.090313.130301

[1] In our context edge orientations are arbitrary and necessary only so as to allow for the definition of an edge inner product in similar terms to that of elementary vector algebra.

algorithm is referred to as the Edge Gram Kernel. A fortunate property of the Edge Gram Kernel is that its elements only depends on one angle and a single pair of distance estimates.

In contrast, the Gram Kernel used in the C-MDS is computed by applying a double-centering transformation [13]–[15] onto the Euclidean distance matrix (EDM), such that each of its elements depends on *all* the distance measurements. In short, the Edge Gram Kernel is, by construction, far more robust to the propagation of measurement errors, which may have devastating effects on the performance of the MDS-based localization algorithm subject to noisy and biased distance estimates [16].

The new MDS formulation is referred to as edge multidimensional scaling (E-MDS), and offers the following advantages over standard MDS solutions [9]. First, it allows a unified processing of angle and ranging information. Second, it leads to a simplified kernel structure that is resilient to the error propagation problem and third it allows to weight the single range information used to infer the sources' locations.[2] Since the proposed algorithm requires the eigen-decomposition of a kernel matrix whose size increases quadratically with the number of nodes in the network it is shown how to incorporate the Nyström approximation [18] to lower the computational complexity to few matrix multiplication only.

The remainder of this article is as follows. After introducing the necessary notation, the classical multidimensional scaling (C-MDS) and the closed form solution to the STRAIN cost function are presented in Section II-A.

Sections II-B and II-C show how to exploit either complete range information or partial range and angle information by means of the C-MDS algorithm. Section III introduces the E-MDS framework and shows how distance and angle information can be processed algebraically (without iteration) and simultaneously. The same section discusses the robustness of the solution to erroneous angle information as well as the possibility to weight the single dissimilarities to cope either with the so called erasure problem, namely partially connected networks.

Section IV explains how to cope with the complexity problem encountered in the E-MDS formulation by means of a spectral approximation of the kernel matrix used in the algorithm. Moreover it is also explained how to extend the range of applicability of the technique to range only multi-source scenario by approximating the angle part of the edge kernel directly from the set of observed distances. Results for both a single and multi-source localization scenarios under both LOS and NLOS are are provided in Section V. Final remarks are offered in Section VI.

II. The MDS Problem over STRAIN

Let $\mathbf{X} \in \mathbb{R}^{N\times\eta}$ denote the matrix containing the set of Cartesian coordinates for N points in an η-dimensional space, then the Euclidean distance between the two generic points $\mathbf{x}_i$ and $\mathbf{x}_j$ is given by

$$d_{ij} = d_{ji} \triangleq \left(\sum_{k=1}^{\eta}(\mathbf{x}_{ik} - \mathbf{x}_{jk})^2\right)^{\frac{1}{2}}. \tag{1}$$

[2]As shown in [17] weighting strategies can substantially improve localization estimates even in presence of scarce ranging information.

Equation (1) can be written as

$$d_{ij}^2 = \mathbf{x}_i \cdot \mathbf{x}_i^{\mathrm{T}} + \mathbf{x}_j \cdot \mathbf{x}_j^{\mathrm{T}} - 2\mathbf{x}_i \cdot \mathbf{x}_j^{\mathrm{T}}, \tag{2}$$

which is the well known *cosine-law* [14].[3]

It follows that the generic squared distance d_{ij}^2 is related to the triangle whose vertices are the points $\mathbf{x}_i$, $\mathbf{x}_j$ and the origin. Also let the squared Euclidean distance matrix (EDM) $\mathbf{D}$ be such that $[\mathbf{D}]_{ij} = d_{ij}^2,\ \forall(i,j) = (\{1,\cdots,N\},\{1,\cdots,N\})$, yielding to

$$\mathcal{D}(\mathbf{X}) = \mathbf{D} = \begin{bmatrix} 0 & d_{1,2}^2 & d_{1,3}^2 & \cdots & d_{1,N}^2 \\ d_{2,1}^2 & 0 & d_{2,3}^2 & \cdots & d_{2,N}^2 \\ d_{3,1}^2 & d_{3,2}^2 & 0 & \cdots & d_{3,N}^2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ d_{N,1}^2 & d_{N,2}^2 & d_{N,3}^2 & \cdots & 0 \end{bmatrix}. \tag{4}$$

While a detailed description of the properties related to $\mathbf{D}$ can be found in [13], [14] it is important to realized that being $\mathbf{D}$ a symmetric matrix with zero diagonal elements, only the $M = \binom{N}{2} = N(N-1)/2$ entries of either its upper or lower triangular part are informative.

Moreover, since any rigid transformation of $\mathbf{X}$ results in the same EDM $\mathbf{D}$, algorithms that rely exclusively on $\mathbf{D}$ to infer $\mathbf{X}$ can only find the solution up to an *isometry* in the embedding space [14]. This uncertainty can be however resolved by the Procrustes transformation detailed in Appendix A which is used to map the solution into the reference system in use [19].

A. The STRAIN Formulation

Let $\mathbb{D}_N(\eta)$ be the space of $N \times N$ distance matrices generated by N points in an η dimensional space and let δ_{ij} to express the dissimilarity [19] between $\mathbf{x}_i$ and $\mathbf{x}_j$. Now consider the squared dissimilarity matrix $\mathbf{\Delta} \in \mathbb{R}^{[N\times N]}$ such that $[\mathbf{\Delta}]_{ij} = \delta_{ij}^2$ and its transformation

$$\mathcal{K}(\mathbf{\Delta},\mathbf{a}) = -\mathbf{J}(\mathbf{a})^{\mathrm{T}} \cdot \mathbf{\Delta} \cdot \mathbf{J}(\mathbf{a}), \tag{5a}$$

$$\mathbf{J}(\mathbf{a}) \triangleq \frac{1}{\sqrt{2}}\left[\mathbf{I} - \mathbf{1}_N \cdot \mathbf{a}^{\mathrm{T}}\right]^{\mathrm{T}}, \tag{5b}$$

where $\mathbf{1}_N \in \mathbb{R}^{N\times 1}$ is a vector whose entries are all one and $\mathbf{a} \in \mathbb{R}^{N\times 1}$ is a *signed distribution* of N terms such that $\sum_{i=1}^{N} a_i = 1$ and whose values influence the *origin* used to represent $\mathbf{X}$ [13].[4]

Schoenberg's embedding theorem for Euclidean distance matrices [21] states that a dissimilarity matrix $\mathbf{\Delta} \in \mathbb{D}_N(\eta)$ if and only if $\mathcal{K}(\mathbf{\Delta},\mathbf{a}) \in \mathbb{S}_N(\eta)$, where $\mathbb{S}_N(\eta)$ is the space of symmetric positive semi-definite (PSD) matrices with rank at most η.[5]

From the above it follows that the metric scaling problem [19] can be seen as the minimization of the distance between

[3]Let θ_{ij}be the angle seen from the origin between the points $\mathbf{x}_i$ and $\mathbf{x}_j$, then equation (2) is equivalent to

$$\cos(\theta_{ij}) = \frac{d_i^2 + d_j^2 - d_{ij}^2}{2d_i d_j}. \tag{3}$$

[4]Notice that equation (5a) is simply the generalization of equation (2), *i.e.* the cosine-law, to all points of $\mathbf{X}$ [20].

[5]Equation (5) is a linear mapping from $\mathbb{D}_N(\eta)$ to $\mathbb{S}_N(\eta)$.

the closed convex set of dissimilarity matrices and the set of symmetric positive matrices [22], yielding to the following optimization problem

$$\begin{aligned} \min_{\mathbf{G}} \quad & \|\mathcal{K}(\mathbf{\Delta}, \mathbf{a}) - \mathbf{G}\|_{\mathrm{F}}^2, \\ \text{s.t.} \quad & \mathbf{\Delta} \in \mathbb{E}_N \\ & \mathbf{G} \in \mathbb{G}_N \end{aligned} \tag{6}$$

where $\mathbb{E}_N$ is the closed convex set of possible dissimilarity matrices and $\mathbb{G}_N$ a subset of $\mathbb{S}_N$.

In the special case of a single dissimilarity matrix, namely $\mathbb{E}_N = \{\mathbf{\Delta}\}$, the optimization problem in equation (6) modifies into

$$\begin{aligned} \min_{\mathbf{G}} \quad & \|\mathcal{K}(\mathbf{\Delta}, \mathbf{a}) - \mathbf{G}\|_{\mathrm{F}}^2, \\ \text{s.t.} \quad & \mathbf{G} \in \mathbb{S}_N(\eta), \end{aligned} \tag{7}$$

where the objective to be minimized is the so called STRAIN cost function.

As shown in [23] and further generalized in [22], an attractive feature of STRAIN is that its global solution can be computed explicitly. Moreover, from Schoenberg's embedding theorem $\mathcal{D}(\mathbf{X}) = \mathbf{\Delta}$ if and only if $\mathcal{K}(\mathbf{\Delta}, \mathbf{a})$ is a Gram matrix (*i.e*, $\mathbf{X} \cdot \mathbf{X}^{\mathrm{T}}$), from which it follows that $\mathbf{X}$ can be obtained by matrix factorization of a PSD matrix.

Let $\mathbf{K_a} = \mathcal{K}(\mathbf{\Delta}, \mathbf{a})$ be a kernel matrix constructed according to equation (5) for a particular value of $\mathbf{a}$, then the coordinate matrix $\mathbf{X}$ can be recovered – up to similarity transformation[6] – by

$$\mathbf{X} = [\mathbf{U_a}]_{1:N,1:\eta} \cdot [\mathbf{\Lambda_a}]_{1:\eta,1:\eta}^{\odot \frac{1}{2}}, \tag{8}$$

where $(\mathbf{\Lambda_a}, \mathbf{U_a})$ is the eigen-pair for $\mathbf{K_a}$ containing the decreasingly ordered eigenvalues and their corresponding eigenvectors respectively, and where $\odot m$ denotes the m-th element-wise power.

Some important generalizations of STRAIN are discussed in [22] while studies on the robustness of STRAIN to perturbations are offered in [24] where it is shown that in the presence of dissimilarities corrupted by zero mean Gaussian noise STRAIN is comparable to a sub-optimal, in the maximum likelihood (ML)-sense, cost function called SSTRESS.

In the particular case of $\mathbf{a} = \mathbf{1}_N/N$, then kernel matrix $\mathbf{G} \in \mathbb{S}_N(\eta)$ can be mapped into $\mathbf{D} \in \mathbb{D}_N(\eta)$ by the following linear operator

$$\mathcal{T}(\mathbf{G}) \triangleq \mathbf{1}_N \cdot \mathrm{diag}(\mathbf{G})^{\mathrm{T}} + \mathrm{diag}(\mathbf{G}) \cdot \mathbf{1}_N^{\mathrm{T}} - 2\mathbf{G}, \tag{9}$$

where $\mathrm{diag}(\cdot)$ is a function that when applied to a matrix returns vector whose elements are identical to the diagonal of the matrix given as argument, while when applied to a vector it returns a diagonal matrix whose nonzero entries are the one in the original vector.

Also notice that the inverse mapping from $\mathbf{D}$ to $\mathbf{G}$ is performed using the operator $\mathcal{K}(\mathbf{\Delta}, \mathbf{a})$ defined in equation (5a) [25].

[6]Similarity transformation includes scaling, rotation and shift and is recovered by Procrustes transformation [19].

B. Complete Distance Information

Assume all the M pairwise distances between the N points in $\mathbf{X}$ as known and consider the *double-centred* Euclidean kernel constructed from equation (5a) setting $\mathbf{a} = \frac{1}{N} \cdot \mathbf{1}_N$ as [14]

$$\mathbf{K} = -\mathbf{J}^{\mathrm{T}} \cdot \mathbf{D} \cdot \mathbf{J}, \tag{10}$$

where

$$\mathbf{J} = \frac{1}{\sqrt{2}} \left[\mathbf{I}_N - \tfrac{1}{N} \cdot \mathbf{1}_N \mathbf{1}_N^{\mathrm{T}} \right]. \tag{11}$$

Then the C-MDS algorithm recovers the estimated configuration of points $\hat{\mathbf{X}}$ from $\mathbf{K}$ using equation (8). This solution represents is exact provided that $\mathbf{D} \in \mathbb{D}_N(\eta)$ and it is known to be the best approximation of $\mathbf{K}$ obtained under the constraint that $\mathrm{rank}(\mathbf{K}) = \eta$ [23]. In practical localization applications, however, one can only obtain a dissimilarity matrix $\tilde{\mathbf{\Delta}}$, corrupted by noise and bias[7], with the obvious consequence that no exact solution is possible. However, under this realistic scenarios a measure of the quality of the approximation obtained by the spectral truncation in equation (8) can be obtained by value of the STRAIN cost function, namely

$$\mathcal{C}_T = \|\hat{\mathbf{X}} \cdot \hat{\mathbf{X}}^{\mathrm{T}} - \mathcal{K}(\tilde{\mathbf{\Delta}}, \mathbf{a})\|_{\mathrm{F}}^2.$$

To gain insight into the effect that corrupted dissimilarities have in the C-MDS algorithm, let a_i be the i-th element of $\mathbf{a}$, then using equation (9) it follows that

$$[\tilde{\mathbf{K}}]_{ij} = -\frac{1}{2}\left([\tilde{\mathbf{D}}]_{ij} - \sum_i a_i [\tilde{\mathbf{D}}]_{ij} - \sum_j a_j [\tilde{\mathbf{D}}]_{ij} + \sum_i \sum_j a_i a_j [\tilde{\mathbf{D}}]_{ij} \right). \tag{12}$$

From the above it is clear that the Euclidean kernel $\tilde{\mathbf{K}}$ used in the C-MDS solution is structured in such a way that each one of its elements depends on the errors of the entire measurement set $\tilde{\mathbf{D}}$ simultaneously, resulting in an error-propagation-prone algorithm [26].

C. Partial Heterogeneous Information

Section II-A showed that through equation (5) the dissimilarity matrix $\mathbf{D}$ can be transformed into the inner product matrix $\mathbf{K}$. Using the definition of inner product between points in $\mathbb{R}^\eta$ then the ij-th element of $\mathbf{K}$ is expressed as

$$[\mathbf{K}_{\mathrm{D}}]_{i,j} \triangleq \langle \boldsymbol{x}_i; \boldsymbol{x}_j \rangle = d_i d_j \cos(\theta_{i,j}). \tag{13}$$

It follows that the matrix $\mathbf{K}_{\mathrm{D}}$ can be written as

$$\mathbf{K}_{\mathrm{D}} = \mathbf{\Theta}_{\mathrm{D}} \otimes \left(\mathbf{d}_N \cdot \mathbf{d}_N^{\mathrm{T}} \right), \tag{14}$$

where $\otimes$ denotes the Hadamard product, $\mathbf{d}_N$ a vector containing the distances from the origin to the N points in $\mathbf{X}$ and

$$\mathbf{\Theta}_{\mathrm{D}} = \begin{bmatrix} 1 & \cdots & \cos(\theta_{1,N}) \\ \vdots & \ddots & \vdots \\ \cos(\theta_{N,1}) & \cdots & 1 \end{bmatrix}, \tag{15}$$

includes all the angles amongst the M edges in $\mathbf{X}$ as seen from the origin.

[7]As discussed later, erasures may also occur.

Using the definition of inner product in equation (14) then $\mathbf{X}$ can be recovered from the kernel matrix $\mathbf{K}_{\mathrm{D}}$ which is obtained from a subset of all possible range and angle measurements. A by-product of this feature is that error propagation is significantly smaller since, as shown in equation (13), the kernel matrix is computed (element-by-element), straight from distance and angle estimates. Also notice that the kernel $\mathbf{K}_D$ only exploits a subset of the entire range and angle information, namely only N out of M range observations and M out of the $\binom{M}{2}$ angles potentially available.

Moreover, to be of any practical use the kernel $\mathbf{K}_{\mathrm{D}}$ needs to be centered at the i-th point of $\mathbf{X}$ from where those distances and angles can be measured. The resulting kernel matrix $\mathbf{K}_{(i)}$ is then equivalent to the one obtained starting from the full set of all mutual distances between the points in $\mathbf{X}$ and computed using equation (5) where the centering vector $\mathbf{a}$ is set to be zero everywhere except in correspondence of the i-th position.[8] Once the new kernel matrix $\mathbf{K}_{(i)}$ is available, then $\mathbf{X}$ can be recovered applying equation (8) followed by Procrustes transformation.

III. Full Heterogeneous Information

Although the kernel matrix $\mathbf{K}_{(i)}$ allows to use angle and distance information jointly, it only includes a subsets of all possible measurements amongst the points in $\mathbf{X}$. In what follows an extension of the C-MDS algorithm that can potentially exploit all mutual distance and angle information measurable between all possible pairs of points in $\mathbf{X}$ is proposed.

Let associate to the network configuration the complete oriented graph $G_{\eta,N}(\mathrm{X},\vec{\mathrm{V}},\mathrm{D})$, where X is the set of N vertices in the η dimensional space and whose coordinate of the i-th vertex is given by the corresponding row of $\mathbf{X}$, $\vec{\mathrm{V}}=\{\vec{\mathrm{v}}_{n,m}\}$ as the set of edges with arbitrary, but unique, orientations (vectors) and D the corresponding weights. For convenience, the orientation of all edges is chosen so that the M elements of $\vec{\mathrm{V}}$ are ordered progressively, *i.e.*, $\vec{\mathrm{V}}=\{\vec{\mathrm{v}}_{1,2},\vec{\mathrm{v}}_{1,3},\cdots,\vec{\mathrm{v}}_{N-1,N}\}$. To simplify the notation, the mutually exclusive pairs of indexes (n,m), with $n<m$, are relabeled by a different number i and let $\mathbf{V}\in\mathbb{R}^{M\times\eta}$ the *edge*-matrix containing the coordinated for all M edges belonging to the graph $G_{\eta,N}(\mathrm{X},\vec{\mathrm{V}},\mathrm{D})$ such that the edge corresponding to the pair of points $(\mathbf{x}_n,\mathbf{x}_m)$ can be denoted simply by,

$$\boldsymbol{v}_i=(\mathbf{x}_m-\mathbf{x}_n)=[(x_{m,1}-x_{n,1}),\cdots,(x_{m,\eta}-x_{n,\eta})]^{\mathrm{T}}. \quad (16)$$

Next, define the dissimilarity metric between the i-th and j-th edges as the inner product,

$$\begin{aligned}[\mathbf{K}_{\mathrm{E}}]_{i,j}&\triangleq\langle\boldsymbol{v}_i;\boldsymbol{v}_j\rangle=\langle(\mathbf{x}_m-\mathbf{x}_n);(\mathbf{x}_q-\mathbf{x}_p)\rangle\\&=d_{n,m}d_{p,q}\cos(\theta_{i,j}).\end{aligned} \quad (17)$$

where the set of all dissimilarity measures $[\mathbf{K}_{\mathrm{E}}]_{i,j}$, corresponding to all pairs of edges in the graph, can be conveniently assembled into the *edge*-kernel

$$\begin{aligned}\mathbf{K}_{\mathrm{E}}&=\langle[\boldsymbol{v}_1,\cdots,\boldsymbol{v}_M];[\boldsymbol{v}_1,\cdots,\boldsymbol{v}_M]^{\mathrm{T}}\rangle\\&=\mathbf{V}\cdot\mathbf{V}^{\mathrm{T}}=\begin{bmatrix}\langle\boldsymbol{v}_1;\boldsymbol{v}_1\rangle&\cdots&\langle\boldsymbol{v}_1;\boldsymbol{v}_M\rangle\\\vdots&\ddots&\vdots\\\langle\boldsymbol{v}_M;\boldsymbol{v}_1\rangle&\cdots&\langle\boldsymbol{v}_M;\boldsymbol{v}_M\rangle\end{bmatrix}.\end{aligned} \quad (18)$$

[8]This is a generalization of the kernel matrix obtained using the full-rank skinny Schoenberg auxiliary matrix $\mathbf{J}=\frac{1}{\sqrt{2}}[-\mathbf{1},\mathbf{I}_{N-1}]^{\mathrm{T}}\in\mathbb{R}^{N\times N-1}$, which centers the reference system in the first point of $\mathbf{X}$ [14].

From equation (17) it follows that the generic element $[\mathbf{K}_{\mathrm{E}}]_{ij}$ is altered only by the errors affecting the i-th and the j-th edge and the corresponding angle. Moreover, differently from the kernel $\mathbf{K}_{\mathrm{D}}$, the edge-kernel $\mathbf{K}_{\mathrm{E}}$ can exploit the entire set of angles and rage measurements potentially available.

The proposed algorithm results directly from the fact that $\mathbf{K}_{\mathrm{E}}$ is a Gram matrix from which it follows that the $\mathbf{V}$ can be recovered from $\mathbf{K}_{\mathrm{E}}$ using equation (8). Namely, the key idea behind the E-MDS algorithm is to solve the cost function in equation (6) for the kernel matrix associated to the graph of $\mathbf{X}$, namely to solve the scaling problem at hand over the set of edges of $G_{\eta,N}(\mathrm{X},\vec{\mathrm{V}},\mathrm{D})$ rather than its set of vertices.

Being interested in the coordinates of the targets, the final step of the localization algorithm is to recover $\mathbf{X}$ from the retrieved vector matrix $\mathbf{V}$. To this end, consider the following system of linear equations, derived directly from equation (16),

$$\mathbf{C}\cdot\mathbf{X}=\mathbf{V}, \quad (19)$$

where the coefficient matrix $\mathbf{C}$ has the block upper-triangular structure shown below

$$\mathbf{C}=\left[\begin{array}{c|c|c|c}\mathbf{1}_{N-1\times1}&-\mathbf{I}_{N-1\times N-1}&&\\\hline\mathbf{0}_{N-2\times1}&\mathbf{1}_{N-2\times1}&-\mathbf{I}_{N-1\times N-1}&\\\hline\ddots&\ddots&\ddots&\\\hline\multicolumn{2}{c|}{\mathbf{0}_{1\times N-2}}&1&-1\end{array}\right]. \quad (20)$$

It is trivial to prove[9] that $\mathrm{rank}(\mathbf{C})=N-1$, such that the system given in equation (19) is under-determined by one equation. In fact, any partition of $N-1$ linearly independent rows of $\mathbf{C}$ has rank $N-1$.

In other words, if it is assumed that the absolute coordinates of a single source (which without loss of generality can be taken to be the first) is known, a solution of the E-MDS problem can be found by considering any of the following reduced systems

$$\left[\begin{array}{c|c}1&\mathbf{0}_{1\times N-1}\\\hline[\mathbf{C}]_{N-1\times N}\end{array}\right]\cdot\left[\begin{array}{c}\mathbf{x}_1\\\hline[\mathbf{X}]_{N-1\times\eta}\end{array}\right]=\left[\begin{array}{c}\mathbf{x}_1\\\hline[\mathbf{V}]_{N-1\times\eta}\end{array}\right], \quad (21)$$

where $[\mathbf{C}]_{N-1\times N}$ denotes an $N-1$ row-partition of $\mathbf{C}$ with rank $N-1$, while $[\mathbf{X}]_{N-1\times\eta}$ and $[\mathbf{V}]_{N-1\times\eta}$ denote the corresponding $N-1$ row-partitions of $\mathbf{X}$ and $\mathbf{V}$, respectively.

Consider the following expression for the inverse of block matrices [27],

$$\left[\begin{array}{c|c}x&\mathbf{y}\\\hline\mathbf{z}&\mathbf{W}\end{array}\right]^{-1}=\left[\begin{array}{c|c}(x-\mathbf{y}\cdot\mathbf{W}^{-1}\cdot\mathbf{z})&-x\cdot\mathbf{y}\cdot(\mathbf{W}-x\cdot\mathbf{z}\cdot\mathbf{y})^{-1}\\\hline-\mathbf{W}^{-1}\cdot\mathbf{z}\cdot(x-\mathbf{y}\cdot\mathbf{W}^{-1}\cdot\mathbf{z})&(\mathbf{W}-x\cdot\mathbf{z}\cdot\mathbf{y})^{-1}\end{array}\right], \quad (22)$$

where x is a scalar, $\mathbf{y}$ and $\mathbf{z}$ row- and column-vectors, and $\mathbf{W}$ an invertible matrix.

It is clear from equation (22) that any of the linear systems described by equation (21) is determined, because $\mathbf{W}$ is an $N-1$ column-partition of $[\mathbf{C}]_{N-1\times N}$, which has rank $N-1$.

[9]Proof: Let $\mathbf{C}_i$ denote the i-th $N-i$ row-partition of $\mathbf{C}$, such that we may write $\mathbf{C}=\left[\begin{array}{c|c|c|c}\mathbf{C}_1^{\mathrm{T}}&\cdots&\mathbf{C}_{N-1}^{\mathrm{T}}\end{array}\right]^{\mathrm{T}}$. Then, simply notice that for all $1\leqslant i<N$, the partition $\mathbf{C}_{i+1}$ is given by the last $N-i-1$ rows of $\mathbf{C}_i$, subtracted by its first row.

A particularly case is the one where $[\mathbf{C}]_{N-1\times N}$ is given by the first $N-1$ rows of $\mathbf{C}$, yielding

$$\left[\begin{array}{c|c} 1 & \mathbf{0}_{1\times N-1} \\ \hline \mathbf{1}_{N-1\times 1} & -\mathbf{I}_{N-1\times N-1} \end{array}\right]\cdot\mathbf{X}_{N\times\eta}=\left[\begin{array}{c} \mathbf{x}_1 \\ \hline \mathbf{V}_{N-1\times\eta} \end{array}\right]. \quad (23)$$

From equation (22) it is found that the coefficient matrix of this reduced linear system is invariable to the inverse operation, so that the solution of the E-MDS becomes simply

$$\mathbf{X}_{N\times\eta}=\left[\begin{array}{c|c} 1 & \mathbf{0}_{1\times N-1} \\ \hline \mathbf{1}_{N-1\times 1} & -\mathbf{I}_{N-1\times N-1} \end{array}\right]\cdot\left[\begin{array}{c} \mathbf{x}_1 \\ \hline \mathbf{V}_{N-1\times\eta} \end{array}\right]. \quad (24)$$

The fact that only $N-1$ rows of $\mathbf{C}$ suffice to solve this reduced version of the E-MDS localization problem is obviously a direct consequence of the fact that there are only $N-1$ linearly independent vectors $\boldsymbol{v}_i$ in V. This in turn implies that turn, means that the algorithm can be formulated over an $N-1$-by-$N-1$ reduced Edge Gram Kernel $\bar{\mathbf{K}}_{\mathrm{E}}=[\mathbf{V}]_{N-1\times\eta}\cdot[\mathbf{V}]^{\mathrm{T}}_{N-1\times\eta}$, where $[\mathbf{V}]_{N-1\times\eta}$ denotes the $N-1$ row-partition of $\mathbf{V}$. Also notice that this solution is equivalent to the scenario with partial heterogeneous information discussed in II-C.

One benefit of the decomposition in equation (18) is its ability to exploit different sources of information, *i.e.* range and angle measurements, in a disjoint manner, contributing to make the overall solution robust to the different sources of errors. This can be better understood by expressing the edge kernel $\mathbf{K}_{\mathrm{E}}$ similarly to equation (14), namely

$$\mathbf{K}_{\mathrm{E}}=\boldsymbol{\Theta}_{\mathrm{E}}\otimes\left(\mathbf{d}_M\cdot\mathbf{d}_M^{\mathrm{T}}\right), \quad (25)$$

where $\boldsymbol{\Theta}_{\mathrm{E}}$ is the *angle*-kernel including all the mutual angles, or better saying its cosines, and $\mathbf{d}_M\in\mathbb{R}^{M\times 1}$ the vector containing the corresponding lengths for all the edges in $G_{\eta,N}(X,\vec{V},D)$.

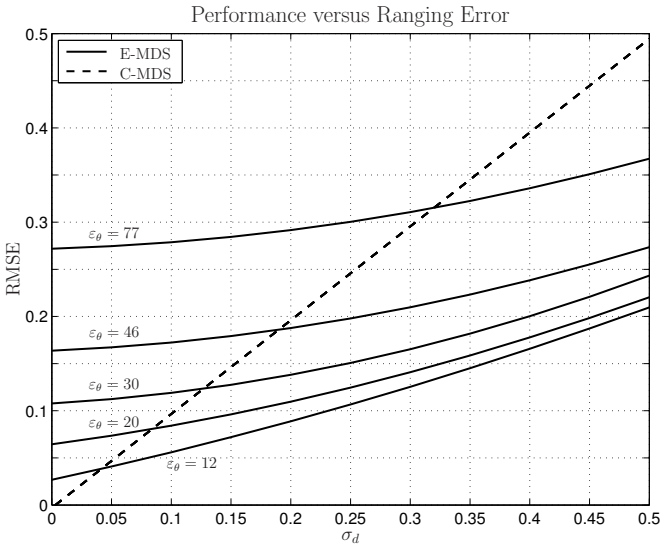
Equation (25) shows how, by using the proposed algorithm, it is possible to exploit the two sources of information in an almost completely decoupled manner by first decomposing $\boldsymbol{\Theta}_{\mathrm{E}}$ into the eigen-pair $(\mathbf{U}_{\boldsymbol{\Theta}_{\mathrm{E}}},\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_{\mathrm{E}}})$ and then finding the estimated edges as

$$\hat{\mathbf{V}}=\mathrm{diag}(\mathbf{d}_M)\cdot[\mathbf{U}_{\boldsymbol{\Theta}_{\mathrm{E}}}]_{1:M,1:\eta}\cdot[\boldsymbol{\Lambda}_{\boldsymbol{\Theta}_{\mathrm{E}}}]^{\odot\frac{1}{2}}_{1:\eta,1:\eta}. \quad (26)$$

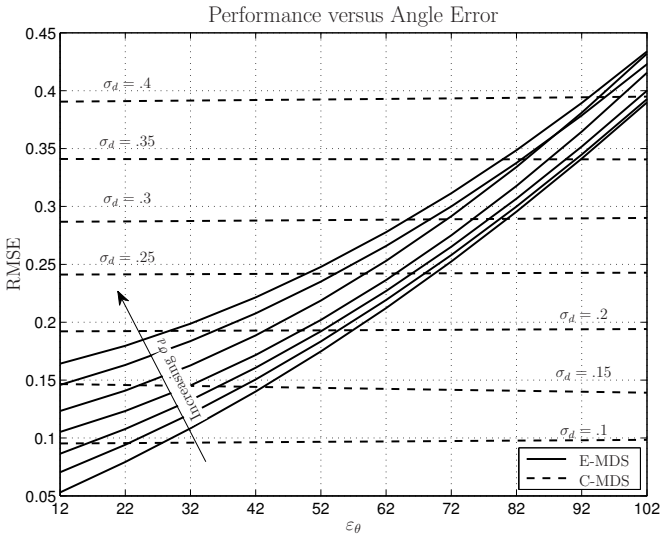
Once $\hat{\mathbf{V}}$ is computed by means of equation (26), then the estimated positions is obtained solving equation (19), namely $\tilde{\mathbf{X}}=\mathbf{C}^{\dagger}\cdot\hat{\mathbf{V}}$, followed by the procrustes transformation detailed in Appendix A. In what follows the performance of the E-MDS technique described above are compared against the C-MDS one, which utilizes only distance information. The simulations are designed to provide insight on how much accuracy can be gained by utilizing angle information or, conversely, how much accuracy is needed from angle estimates in order to reap better accuracies in comparison to using only distance measurements.

In each realization, a fully connected random network of sensors Gaussianly distributed in the 2-dimensional Euclidean space, with statistical center at the origin and a unitary spread (variance) is generated. All metric units are normalized and, therefore, shall be omitted henceforth.

Range estimations $\tilde{d}$, *i.e.*, the noisy measurements of the distance between each pair of sensors, are modeled as



(a) RMSE as a function of σ_d.



(b) RMSE as a function of ε_θ.

Fig. 1. Performance of metric and E-MDS as a function of σ_d and ε_θ.

Gaussian-distributed random variables with the mean given by the true distance d and a standard deviation σ_d related to the ranging error affecting its measurement, *i.e.*,

$$p_{\mathrm{Gauss}}(r;d,\sigma_d)\triangleq\frac{1}{\sigma\sqrt{2\pi}}\cdot\exp\left(\frac{-(r-d)^2}{2\sigma_d^2}\right). \quad (27)$$

The complete EDM sample corresponding to these Gaussian-distributed distance estimates are then fed to the C-MDS localization algorithm described in Section II-B.

In turn, angle estimation errors are modeled as random processes distributed according to the Tikhonov probability density function (pdf) [28]

$$p_{\mathrm{T}}(\theta;\rho)=\frac{1}{2\pi I_0(\alpha)}\cdot e^{\rho\cos(\theta)},\ \theta\in[-\pi,\pi],\ \rho\geqslant 0. \quad (28)$$

The parameter ρ controls the shape of this pdf, such that $p_{\mathrm{T}}(\theta;\rho)$ tends to a uniform distribution for $\rho\to 0$, and to a Dirac delta at 0 when $\rho\to\infty$. For a given ρ, we shall define the angle error ε_θ as the limiting angle that encloses 90% of

the area below $p_{\mathrm{T}}(\theta;\rho)$, *i.e.*,

$$\varepsilon_\theta = \theta_{\mathrm{L}} \left| \int_{-\theta_{\mathrm{L}}}^{\theta_{\mathrm{L}}} p_{\mathrm{T}}(\theta;\rho)\, d\theta = 0.9. \right. \tag{29}$$

The range and angle estimates perturbed according to the Gaussian and Tikhonov distributions are then fed to the E-MDS localization algorithm described above.

The performance of the algorithms are measured by the Frobenius norm of the difference between the estimate coordinate vector $\hat{\mathbf{X}}$ and the corresponding true coordinate vector $\mathbf{X}$, namely the root mean square error (RMSE), *i.e.*,

$$\mathrm{RMSE} \triangleq \tfrac{1}{\sqrt{N}} \cdot \|\mathbf{X} - \hat{\mathbf{X}}\|_2. \tag{30}$$

The algorithms are compared in Figure 1 and Figure 2. First, the estimation accuracies attained by both algorithms as functions of the ranging error are shown in Figure 1(a). In the case of the E-MDS algorithm, the curves are parameterized by the angle error ε_θ.

The plots show that the localization accuracy of the E-MDS algorithm grows (almost linearly) with σ_d at a lower pace than that of the C-MDS in the interval of interest (of moderate ranging errors). In other words, it is found that angle information, even when not highly accurate, can significantly improve the localization accuracy.[10]

For example, in a scenario with $\sigma_d = 0.2$ and $\varepsilon_\theta \approx 20°$, the E-MDS algorithm exhibits the same performed attained by the C-MDS when the range estimation errors are inferior to 0.1.

The fact that angle information, even if imperfect, can be utilized to increase the robustness versus ranging errors can be further appreciated in Figure 1(b), where the localization accuracy of the E-MDS algorithm as a function of AOA estimation errors and parameterized by σ_d – varying from 0.1 to 0.4 – are compared against the accuracy of the corresponding C-MDS solution. As the plot illustrates, although the accuracies of both algorithms are affected by an increase on the range estimation error, the relative performance degradation suffered by the E-MDS solution (for any given underlying ε_θ) is significantly lower than the one experienced by its classical counterpart.

Also notice that the mapping defined in equation (19) allows to *weight* the rows of $\mathbf{C}$ and consequently the single entries of $\mathbf{D}$. For instance, by weighting to zero the rows of $\mathbf{C}$ corresponding to the missing entries in $\mathbf{D}$ then it is possible to make the E-MDS algorithm robust to the distance erasures problem, which is know to be an extremely limiting factor for the applicability of the C-MDS algorithm [9]. This is illustrated in Figure 2 where it is shown that, for distance information corrupted by $\sigma_d = 0.05$ and different values of ε_θ the performance of the E-MDS are almost independent on the level of completeness of $\mathbf{D}$. This robustness to of the algorithm versus missing measurements can be understood looking at the single elements defined in equation (17) as well as the structure of the Edge kernel $\mathbf{K}_{\mathrm{E}}$. Indeed while equation (18) shows how each single entry of the kernel only depends on two distances and the corresponding mutual angle it also illustrates

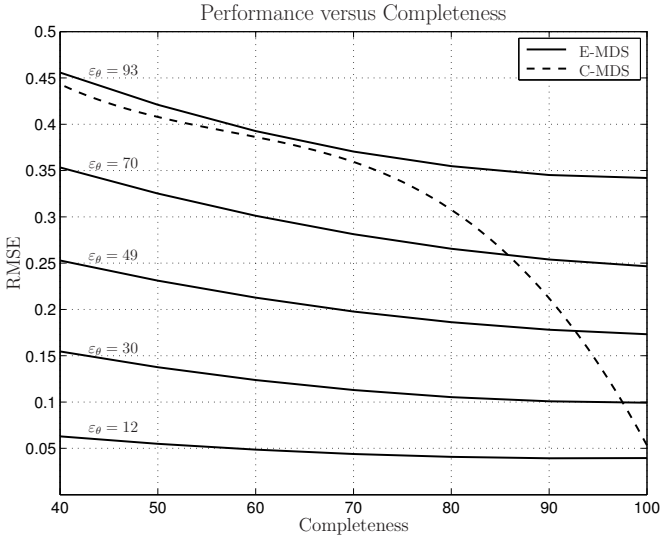


Fig. 2. Robustness of the E-MDS versus incomplete information.

how the same range information is used *redundantly* in $M-1$ combinations.

Moreover, another strong point of the E-MDS solution compared to the standard C-MDS, is that eventual knowledge on the *confidence* of the single range measurements could be incorporated in the estimation process through the matrix $\mathbf{C}$, for instance relying on techniques such as the one proposed in [17], [26].

IV. The E-MDS in Localization Scenarios

In the form presented in Section III the E-MDS algorithm still suffers the following limitations:

- *increased complexity to eigen-decompose* $\mathbf{K}_{\mathrm{E}}$,
- *all the angles in* $\mathbf{\Theta}_{\mathrm{E}}$ *are required.*

Indeed, being $\mathbf{K}_{\mathrm{E}} \in \mathbb{R}^{M\times M}$, where $M = \binom{N}{2}$, one bottleneck of the algorithm is represented by the complexity required to eigen-decompose $\mathbf{K}_{\mathrm{E}}$, which, in presence of networks of sufficiently large size could become unbearable.

Fortunately, however, depending on the kind of scenario at hand, different kind of solutions can be exploited to drastically reduce the aforementioned computational requirements. With reference to the problem at hand we will consider the following two cases, complete and structurally incomplete information in the angle kernel $\mathbf{\Theta}_{\mathrm{E}}$ in equation (25). Specifically, in the former case, although different kind of algorithms could be used to alleviate the computation associated with the eigen-decomposition of $\mathbf{\Theta}_{\mathrm{E}}$ [27], in [9] it was shown that in dynamic scenarios its eigenspace can be tracked at a computational cost that grows only linearly with $\sqrt{N}$. Nevertheless, while in this scenario the computational complexity does not pose a significant restriction on the applicability of the algorithm, a more stringent requirement comes from the underline assumption that, although with low accuracy, all the M mutual angles amongst the edges in $\vec{\mathrm{V}}$ need to be known.[11]

Fortunately, however, in many scenarios of interest, *e.g.* cellular network like scenarios, it can be expected that only

[10]This is in line with the theory which states that the Fisher information provided by independent measurement is the sum of the information from single observations [29].

[11]Notice that the incompleteness discussed in Section III refers to the distance information, which in light of equation (25) it is equivalent to a *structured* incompleteness in $\mathbf{\Theta}_{\mathrm{E}}$.

a subset of nodes, *e.g.* the base stations, are equipped with the hardware necessary to provide the AOA measurements. Under this assumption then the completeness of $\mathbf{\Theta}_{\mathrm{E}}$ is limited to those rows corresponding to the edges in which at least one of the two vertices is an anchor node.

In presence of this *well-structured* incompleteness of $\mathbf{\Theta}_{\mathrm{E}}$, then it becomes possible to exploit a numerical approximation to the eigenfunction problem used in the context of kernel based predictors which goes under the name of Nyström approximation [18] to make the computational complexity required to recover $\hat{\mathbf{V}}$ substantially lower.

Indeed using the Nyström algorithm it is possible to approximate the eigen-problem associated to $\mathbf{K}_{\mathrm{E}}$ with one of size $\eta = \mathrm{rank}(\mathbf{K}_{\mathrm{E}})$, and then to expand back the result [18]. Let B_{A} be a number of rows selected at random from the kernel matrix[12], $B_{\mathrm{T}} = M - B_{\mathrm{A}}$ the remaining edges in $G_{\eta,N}(\mathrm{X}, \vec{\mathrm{V}}, \mathrm{D})$ and without loss generality let's assume that those are permuted to the first B_{A} position of the the matrix, yielding

$$\mathbf{K}_{\mathrm{E}} \approx \left[\begin{array}{c|c} \mathbf{A}_{B_{\mathrm{A}}\times B_{\mathrm{A}}} & \mathbf{T}_{B_{\mathrm{A}}\times B_{\mathrm{T}}} \\ \hline \mathbf{T}^{\mathrm{T}}_{B_{\mathrm{T}}\times B_{\mathrm{A}}} & \mathbf{T}^{T}_{B_{\mathrm{T}}\times B_{\mathrm{A}}} \cdot \mathbf{A}^{-1}_{B_{\mathrm{A}}\times B_{\mathrm{A}}} \cdot \mathbf{T}_{B_{\mathrm{A}}\times B_{\mathrm{T}}} \end{array}\right]. \quad (31)$$

Let $\mathbf{V} = \left[\mathbf{V}_{\mathrm{A}}^{\mathrm{T}}, \mathbf{V}_{\mathrm{T}}^{\mathrm{T}}\right]^{\mathrm{T}}$ and $\mathbf{A} = \mathbf{U}_A \cdot \mathbf{\Lambda}_A \cdot \mathbf{U}_A^{\mathrm{T}}$, then using equation (8) it follows that $\mathbf{V}_{\mathrm{A}} = [\mathbf{U}_{\mathrm{A}}]_{B_{\mathrm{A}}\times\eta} \cdot [\mathbf{\Lambda}_{\mathrm{A}}]^{\odot\frac{1}{2}}_{1:\eta}$, and since $\mathbf{T} = \mathbf{V}_{\mathrm{A}} \cdot \mathbf{V}_{\mathrm{T}}^{\mathrm{T}}$ then

$$\mathbf{V}_{\mathrm{T}} = \left(\mathbf{V}_{\mathrm{A}}^{-\mathrm{T}} \cdot \mathbf{T}\right)^{\mathrm{T}} = \left([\mathbf{\Lambda}_{\mathrm{A}}]^{\odot\left(-\frac{1}{2}\right)}_{\eta\times\eta} \cdot [\mathbf{U}_{\mathrm{A}}]^{\mathrm{T}}_{B_A\times\eta} \cdot \mathbf{T}\right)^{\mathrm{T}}. \quad (32)$$

Hence the matrix $\mathbf{V_T}$ containing the edges referred to the targets can be recovered from the eigen-decomposition of $\mathbf{A}$ only. Moreover, begin $\mathbf{A}$ the edge kernel associated to the anchors, then its eigen-space can be fed to the algorithm as an a priori information.

In light of the discussion above it becomes evident that this approach is particularly convenient in all those cases, such as the multitarget localization scenario here considered in which $B_{\mathrm{A}} \ll B_{\mathrm{T}}$. Indeed, due to the combinatorial relationship existing between the number of nodes and edges in a graph, when $N_{\mathrm{T}} > N_{\mathrm{A}}$ then $B_{\mathrm{T}} \gg B_{\mathrm{A}}$. It follows that the Nyström approximation of the kernel $\mathbf{K}_{\mathrm{E}}$ only relies on the first B_{A} rows of $\mathbf{K}_{\mathrm{E}}$. More importantly the edges associated to the targets can be computed directly by means of equation (32), namely without involving any decomposition.

A further implication of the Nyström method can be understood by expressing the edge kernel $\mathbf{K}_{\mathrm{E}}$ similarly to equation (25). Indeed, using equation (31) an approximation of $\mathbf{K}_{\mathrm{E}}$ can be constructed on the basis of

$$\mathbf{A} = \mathbf{\Theta}_{\mathrm{A}} \otimes \left(\mathbf{d}_{B_{\mathrm{A}}} \cdot \mathbf{d}^{\mathrm{T}}_{B_{\mathrm{A}}}\right), \quad (33)$$

$$\mathbf{T} = \mathbf{\Theta}_{\mathrm{T}} \otimes \left(\mathbf{d}_{B_{\mathrm{T}}} \cdot \mathbf{d}^{\mathrm{T}}_{B_{\mathrm{T}}}\right), \quad (34)$$

where $\mathbf{\Theta}_{\mathrm{A}}$ and $\mathbf{\Theta}_{\mathrm{T}}$ are respectively the matrices including all the anchor-to-anchor and anchor-to-target angles and $\mathbf{d}_{B_{\mathrm{A}}} \in \mathbb{R}^{B_{\mathrm{A}}\times 1}$ and $\mathbf{d}_{B_{\mathrm{T}}} \in \mathbb{R}^{B_{\mathrm{T}}\times 1}$ the vectors with the lengths for the corresponding edges.

[12]Provided that the entries in $\mathbf{A}$ and $\mathbf{T}$ are correct, then the Nyström approximation of $\mathbf{K}_{\mathrm{E}}$ is perfect [18].

Algorithm 1 E-MDS localization

1: Get $\tilde{\mathbf{d}}_M \in \mathbb{R}^{M\times 1}$ and $\tilde{\mathbf{\Theta}}_{\mathrm{E}} \in \mathbb{R}^{M\times M}$;
2: **if** $\tilde{\mathbf{\Theta}}_{\mathrm{E}}$ is complete **then**
3: Factorize $\mathbf{\Theta}_{\mathrm{E}}$ into $(\mathbf{U}_{\mathbf{\Theta}_{\mathrm{E}}}\, \mathbf{\Lambda}_{\mathbf{\Theta}_{\mathrm{E}}})$;
4: Solve $\hat{\mathbf{V}} \leftarrow \mathrm{diag}(\mathbf{d}_M) \cdot [\mathbf{U}_{\mathbf{\Theta}_{\mathrm{E}}}]_{1:M,1:\eta} \cdot [\mathbf{\Lambda}_{\mathbf{\Theta}_{\mathrm{E}}}]^{\odot\frac{1}{2}}_{1:\eta,1:\eta}$;
5: **else**
6: Get $(\mathbf{U}_{\mathbf{\Theta}_{\mathrm{A}}}, \mathbf{\Lambda}_{\mathbf{\Theta}_{\mathrm{A}}})$ and $\mathbf{\Theta}_{\mathrm{T}}$;
7: Solve $\hat{\mathbf{V}}_{\mathrm{T}} \leftarrow \mathrm{diag}(\mathbf{d}_{B_{\mathrm{T}}}) \cdot \left([\mathbf{\Lambda}_{\mathbf{\Theta}_{\mathrm{A}}}]^{\odot\left(-\frac{1}{2}\right)}_{\eta\times\eta} \cdot [\mathbf{U}_{\mathbf{\Theta}_{\mathrm{A}}}]^{\mathrm{T}}_{B_A\times\eta} \cdot \mathbf{\Theta}_{\mathrm{T}}\right)^{\mathrm{T}}$;
8: Compute $\hat{\mathbf{V}} \leftarrow [\mathbf{V}_{\mathrm{A}}^{\mathrm{T}}, \hat{\mathbf{V}}_{\mathrm{T}}^{\mathrm{T}}]^{\mathrm{T}}$;
9: **end if**
10: Map the edges in $\hat{\mathbf{V}}$ into points: $\tilde{\mathbf{X}} \leftarrow \mathbf{C}^{\dagger} \cdot \hat{\mathbf{V}}$;
11: Use $\tilde{\mathbf{X}}$ to compute the Procrustes coefficients $\{\alpha, \mathbf{F}_{\mathrm{R}}, c_{\mathrm{D}}\}$
12: Solve $\hat{\mathbf{X}} \leftarrow \alpha \cdot \mathbf{Y} \cdot \mathbf{F}_{\mathrm{R}} + \mathbf{c}_{\mathrm{D}} \otimes \mathbf{1}_N$.

Since $\mathbf{d}_{B_A} = [\mathbf{d}_{B_A}, \mathbf{d}_{B_T}]$ then it is evident that the Nyström approximation retains all range information in $\mathbf{d}_M$ and only discards part of the information in $\mathbf{\Theta}_{\mathrm{E}}$ by neglecting the mutual angles between the targets.[13]

Therefore, similarly to equation (26), the case of partial angle information can be rewritten as the solution of the edges orientation problem rescaled appropriately, namely

$$\begin{aligned} \mathbf{V}_{\mathrm{T}} &= \mathrm{diag}(\mathbf{d}_{B_{\mathrm{T}}}) \cdot \left(\mathbf{\Theta}_{\mathrm{A}}^{-\mathrm{T}} \cdot \mathbf{\Theta}_{\mathrm{T}}\right)^{\mathrm{T}} \\ &= \mathrm{diag}(\mathbf{d}_{B_{\mathrm{T}}}) \cdot \left([\mathbf{\Lambda}_{\mathbf{\Theta}_{\mathrm{A}}}]^{\odot\left(-\frac{1}{2}\right)}_{\eta\times\eta} \cdot [\mathbf{U}_{\mathbf{\Theta}_{\mathrm{A}}}]^{\mathrm{T}}_{B_A\times\eta} \cdot \mathbf{\Theta}_{\mathrm{T}}\right)^{\mathrm{T}}, \end{aligned} \quad (35)$$

where $\mathbf{\Theta}_{\mathrm{A}} = \mathbf{U}_{\mathbf{\Theta}_{\mathrm{A}}} \cdot \mathbf{\Lambda}_{\mathbf{\Theta}_{\mathrm{A}}} \cdot \mathbf{U}^{\mathrm{T}}_{\mathbf{\Theta}_{\mathrm{A}}}$.

The E-MDS solution applied to heterogeneous scenarios is summarized in Algorithm 1, from which it evident how, at least of the case of partial information, the most computationally demanding part is the computation of $\hat{\mathbf{V}}$ which is obtained by a mere matrix multiplication.

A. The E-MDS with Angle Completion

Although Section III already dealt with the problem of eventual erasures in $\mathbf{d}_M$, the E-MDS approach still requires that all, or using the Nyström approximation, that at least the angles in the first B_A rows of $\mathbf{\Theta}_{\mathrm{E}}$ are known.

Fortunately, however, in light of the robustness of the E-MDS method to erroneous angle information [30] it is possible to either complete or even replace the angle matrix $\mathbf{\Theta}_{\mathrm{E}}$ by

$$\hat{\mathbf{\Theta}}_{\mathrm{E}} = \left(\hat{\mathbf{V}}_{\mathrm{E}} \cdot \hat{\mathbf{V}}_{\mathrm{E}}^{\mathrm{T}}\right) \oslash \left(\mathbf{d}_{\mathrm{E}} \cdot \mathbf{d}_{\mathrm{E}}^{\mathrm{T}}\right), \quad (36)$$

[13]Similarly to [9], generic scenarios in which a subset of the target nodes are able to measure some of the angles to their neighbors can be dealt with by combining the solution described in Section III together with the Nyström approximation.

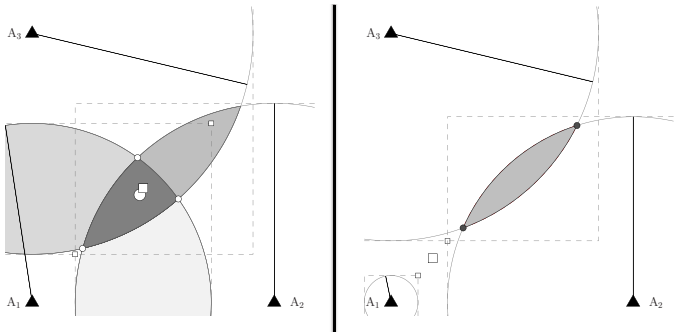


Fig. 3. Feasibility region Φ defined accordingly to equation (39) and (41) in presence of positive biases (left) and positive and negative biases (right).

where $\hat{\mathbf{V}}_{\mathrm{E}} = \mathcal{V}(\hat{\mathbf{X}})$ is the matrix of edges constructed from an initial estimates of $\hat{\mathbf{X}}$, $\oslash$ is the inverse of $\otimes$ and $\mathbf{d}_{\mathrm{E}} \in \mathbb{R}_{+}^{M\times 1}$ is defined as

$$\mathbf{d}_{\mathrm{E}} = \left(\mathrm{diag}\left(\hat{\mathbf{V}}_{\mathrm{E}} \cdot \hat{\mathbf{V}}_{\mathrm{E}}^{\mathrm{T}}\right)\right)^{\odot \frac{1}{2}}. \tag{37}$$

Once $\hat{\mathbf{\Theta}}_{\mathrm{E}}$ is computed, then the new edge kernel is obtained by

$$\hat{\mathbf{K}}_{\mathrm{E}} = \hat{\mathbf{\Theta}}_{\mathrm{E}} \otimes \left(\mathbf{d}_M \cdot \mathbf{d}_M^{\mathrm{T}}\right). \tag{38}$$

Unlike intuition may lead us to believe the angle information constructed from distances is in fact extrinsic information, and therefore, from an Information Theoretical stand point fundamentally contributes to the potential reduction of localization errors. To elaborate, notice that angles, or better saying, their cosines in $\mathbf{\Theta}_{\mathrm{E}}$, are computed from ratios of distances (see equation (3)), which are affected by independent and random errors.[14] In view of such independence between distances and their derived angles, allied with the fact that the addition of any type of extrinsic (independent) information to the estimator leads to smaller Cramer Rao bounds [29], it follows that in fact AOA obtained based on TOA measurement does provide additional useful information to improve positioning accuracy.

Also notice that although reconstructed angles are independent from the distances from which they are calculated, they are not fully uncorrelated amongst themselves. Indeed, each of the random variables θ_{ij} is in fact independent form the the corresponding distances (d_i, d_j). However, the pairs of variates $(\theta_{ij}, \theta_{ik})$ do exhibit correlation between them. Such correlation is not present, of course, if (θ_{ij} and θ_{ik}) are independently measured as assumed in Section III. Secondly, although no correlation exists amongst θ_{ij}, d_i and d_j, if θ_{ij} is reconstructed form d_i and d_j (as well as d_{ij}), the variance of θ_{ij} is determined by the variances of the noisy distances from which it is determined. In contrast, when θ_{ij}'s are independently measured, their variances depend only on the quality of the corresponding estimator. Therefore, even when angles cannot be independently measured, the utilization of range-only in the E-MDS is still advantageous.

However, the problem of finding an appropriate estimate for $\hat{\mathbf{\Theta}}_{\mathrm{E}}$, or equivalently $\hat{\mathbf{V}}_{\mathrm{E}}$, still remains. One method to compute $\hat{\mathbf{\Theta}}_{\mathrm{E}}$ on the basis of the set of anchor-to-target distance measurements only is described in the following.[15]

Let $\{d_{1,j}, \ldots, d_{N_A,j}\}$ be the set of distance measurements between the j-th target and the anchor-to-target measurements, then it is well known that $[\hat{\mathbf{X}}_C]_j$ can be found jointly solving a set of $i = \{1, \ldots, B_A\}$ constraints [33].

In particular, in the case of planar configurations ($\eta = 2$), the extension to the case of 3-D scenarios is straightforward, and denoting with $\hat{\mathbf{x}}_j = [x_j, y_j]$ the j-th target's coordinates, then the relation with respect to the i-th anchor can be expressed by

$$\| \mathbf{x}_i - \mathbf{x}_j \| = d_{i,j} \Rightarrow (x_i - x_j)^2 + (y_i - y_j)^2 = d_{i,j}. \tag{39}$$

Under noisy range measurements equation (39) does not hold. In particular, in presence of NLOS conditions, then the equality in equation (39) modifies into the following inequality [33]

$$\| \mathbf{x}_i - \mathbf{x}_j \| \leqslant d_{i,j} \Rightarrow (x_i - x_j)^2 + (y_i - y_j)^2 \leqslant d_{i,j}. \tag{40}$$

It is also known that each one of the i-th circular constraints above can be relaxed and expressed by a set of η linear constraints [33] as

$$x_i - x \leqslant d_{i,1} \quad \text{and} \quad -x_i + x \leqslant d_{i,1}, \tag{41}$$

$$y_i - y \leqslant d_{i,1} \quad \text{and} \quad -y_i + y \leqslant d_{i,1}. \tag{42}$$

In particular, let $\phi^{(i)}$ be the area defined by the i-th constraint relative to the j-th target, than the intersection the intersection of the B_A constraints defines a *feasibility* region in the space where the target is ensured to be, namely

$$\mathbf{x}_j \in \Phi = \bigcap_{i=1}^{N_A} \phi^{(i)}. \tag{43}$$

A representation of the *feasibility* spaces obtained from the B_A inequalities given in equation (39) or the linear constraints defined in equation (41) is provided in Figure 3 from which it is evident that the region obtained by the union of all linear constraints is only an upper bound to one obtained using the circular constraints.[16]

In absence of any additional information about the range measurements, *e.g* statistical models, then each point in the feasibility region Φ can be assumed as equally likely, implying that the most conservative estimate of $\hat{\mathbf{x}}_j$ can be obtained as the center of mass of the region itself. Specifically for the circular constraints such a point can be found by intersecting the radial lines defined by the pair or inequalities [34]. Differently the feasibility region defined by the linear constrains can be fully characterized by a *lower* and an *upper* point whose coordinates are [33]

$$\mathbf{x}_l = \left[\min_i\{x_i + d_{i,j}\}, \min_i\{y_i + d_{i,j}\}\right], \tag{44}$$

$$\mathbf{x}_u = \left[\max_i\{x_i - d_{i,j}\}, \max_i\{y_i - d_{i,j}\}\right]. \tag{45}$$

[14]It has been shown [31] that the variables Y and $1/Y$ are independent, which in turn implies that the ratio Z = X/Y of two independent variates X and Y is such that Z is independent of both X and Y.

[15]That does not preclude the utilization of any other more sophisticated, if more complex method to compute centroids from feasibility regions. One example of an alternative (substantially more computationally demanding, but also more accurate) method that could be used can be found, for instance, in [32].

[16]Also notice that while the system of linear constraints can always be applied, the system of circular constraints can be solved only in presence of perfect or positively biased range measurements [33].

It follows that the center of mass of for the linear constraints can be found by simply averaging the aforementioned points.

V. Results

In order to evaluate the performance of the E-MDS algorithm described in Section IV, let us consider a scenario consisting of 4 anchor nodes placed at the edges of a 2×2 square in a $\eta = 2$ dimensional space which will also define our surveillance region. The noisy range measurements are modeled as a Gaussian-distributed random variables with mean given by the true distance and a standard deviation σ_d, *i.e.* $d_{ij} \sim p_{Gauss}(r; d, \sigma_d)$, while the angle error is modeled using the Tikhonov model described in Section III. NLOS conditions are simulated by adding to the mean of the aforementioned Gaussian random variable a sample from an random variable uniformly distributed between $0 - b_d$, *i.e.* $\mathcal{U}(0, b_d)$.

To begin with, a source localization in a range-only scenario is considered in Figure 4 where the E-MDS-based solution using the feasibility constraints defined in Section IV-A is compared under both LOS and NLOS conditions versus state-of-the-art algorithms.

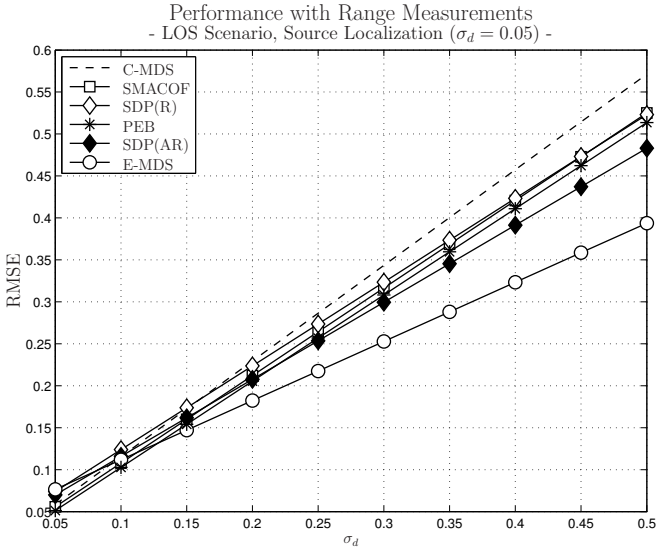
From Figure 4 it can be appreciated that in fact the approach of obtaining angle estimates from given distance measurements, followed by localization via E-MDS, outperforms not only the classic (range-only) MDS, but also other widely used algorithms such as SMACOF and a semi-definite programming (SDP) formulation of the problem [35].

Also Figure 4 compares the proposed solution versus the position error bound (PEB) for range-only scenario, namely the CRLB in LOS and its generalization in NLOS case [4]. In particular, the apparently unreasonable result in Figure 4(a) that sees the E-MDS outperforming the bound can be explained by recalling that, even in presence of range-only scenario the angle kernel $\hat{\mathbf{\Theta}}_{\mathrm{E}}$ from the centroids provides extrinsic information to improve the estimation of $\hat{\mathbf{X}}$.[17] To corroborate on this notice that also the SDP formulation exploiting heterogenous information proposed in [35], namely the SDP(AR) solution, when fed with angles computed from the centroids also beats the bound.

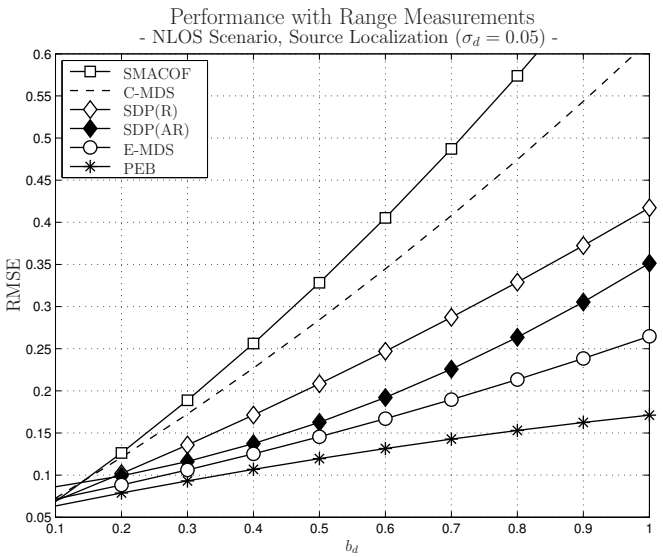
Moreover, when distances are measured almost without error, the correspondingly constructed angles from the centroids become almost deterministic, thus providing no additional statistical information from which to obtain better results. This is precisely why, in Figure 4(a), it can be seen that the performance of the proposed S-MDS algorithm reduces to the same as those of all other methods when σ_d is small.

Regarding the improved performance of E-MDS in Figure 4(b), one reason for such include, as discussed in Section III, the fact that neither the operations of angle calculation from distances, nor the construction of the S-MDS kernel, require all distances (per angle or per kernel element), such that error propagation is drastically reduced. The other reason discussed in Section IV-A stems from the deeper implication of the joint utilization of angle and distance information onto the Fisher Information Matrix associated with the problem.

[17] A more *accurate* bound bound would need to take into consideration this information. Also notice that this happens in the LOS case where the centroids are more *informative*.



(a) Performance in LOS. RMSE as a function of σ_d.



(b) Performance in NLOS. RMSE as a function of the bias b_d.

Fig. 4. Performance of E-MDS algorithm applied to a range only source localization scenario. Results as function of σ_d in the LOS case and b_d for the NLOS case.

Figure 5 offers an additional set of results for the source localization setting in which the E-MDS algorithm is compared versus the SDP(AR) in a scenario in which both range and angle information are assumed to be measured independently from each other. Similarly to [35], angles are assumed to be measured at the sensors accordingly to their local reference system which is assumed to be known only at the anchor nodes. Anyhow, by subtracting the relative measurements at each sensor with respect to other sensors, namely assuming an angle difference of arrival (ADOA) scenario, the dependence from the local reference systems is removed from the problem.

Moreover, as explained in [35] the SDP(AR) considers the Law of Cosines applied to all triplets of points in the scenario, namely a typical ADOA scenario, while the angle kernel $\mathbf{\Theta}_{\mathrm{E}}$

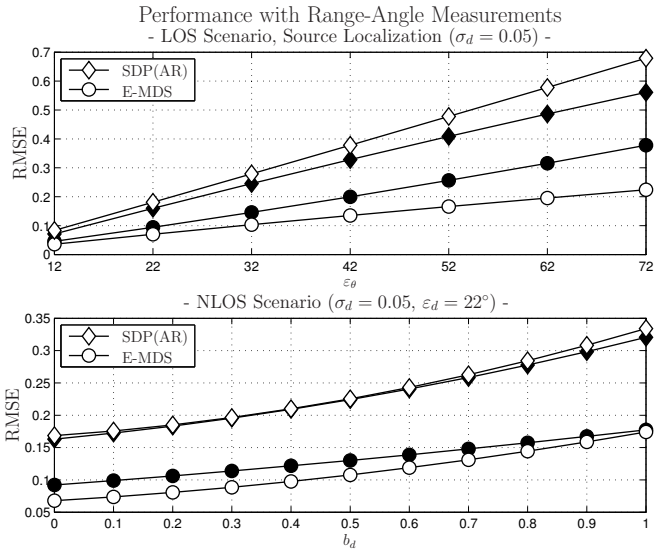
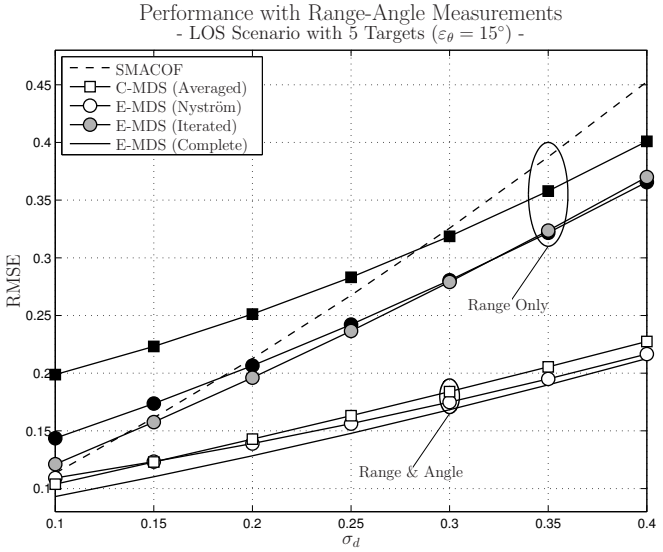
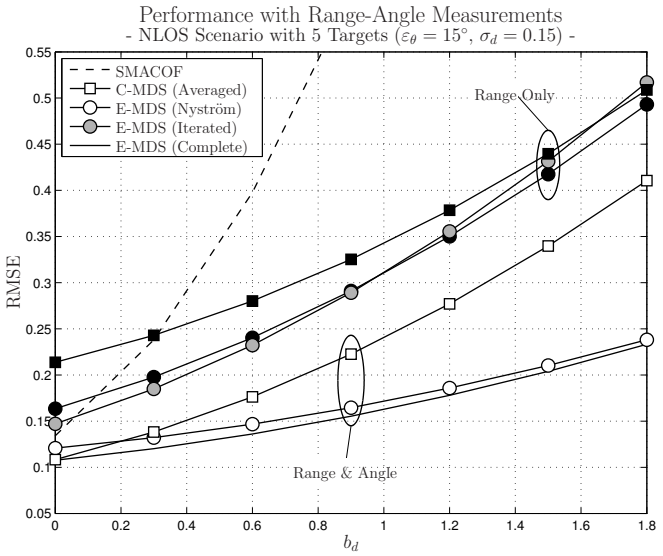


Fig. 5. Performance of E-MDS algorithm applied to a range-angle source localization scenario. Results as function of ε_d for the LOS case and b_d for the NLOS case.



(a) Performance in LOS. RMSE as a function of σ_d.



(b) Performance in NLOS. RMSE as a function of the bias b_d.

Fig. 6. Performance of E-MDS algorithm applied to a multitarget scenario. Results as function of σ_d and b_d, for range only and range-bearing measurements.

extends it to *all pairs* of edges in $G_{\eta,N}(\mathrm{X}, \vec{\mathrm{V}}, \mathrm{D})$, namely it exploits more information than the SDP(AR) algorithm.[18]

The second scenario considered consists of a multi target localization scenario in which 5 targets are randomly displaced within the surveillance region. As discussed in Section III and IV, the E-MDS framework allows to handle multitarget dynamic scenarios in which either distance or distance and AOA information is available. In what follows the performance of the E-MDS technique is compared against the SMACOF algorithm [19] and the C-MDS based method described in Section II-C using both range only and range and angle measurements.

The simulations are designed to provide insight on how much accuracy can be gained by utilizing angle information and how robust the edge-kernel formulation is compared to the other alternatives discussed in Section II.

As discussed in Section IV, although the E-MDS algorithm allows to include range and angle measurements from each pair of sensors, in practice it can not be expected that every device will be able to collect these kind of information. For this reason below we will restrict ourselves to a cellular network-like scenario in which only a limited number (N_A) of nodes, the anchor nodes, can measure angle and range with all the users in their surroundings.

The results are shown in Figure 6 for a LOS and NLOS conditions. The plots reveal that the E-MDS solution implemented as discussed in Section IV has better accuracy than both SMACOF algorithm[19] and the MDS-based method proposed in Section II-C averaged over all N_A anchor nodes. As Figure 6(b) shows, the advantage of the E-MDS solution becomes particularly evident in presence of NLOS measurements. Moreover, similarly to iterated Bayesian solutions, *e.g.* iterated Kalman filter [29], in the case of range only observations an additional improvement in performance can be achieved by updating $\hat{\mathbf{\Theta}}_C$ on the basis of the solution obtained by the E-MDS and running the algorithm once more time over the observed range measurements.

VI. Conclusions

We proposed a comprehensive algorithm to handle the problem of simultaneously localize a large number of targets in LOS-NLOS with no a priori measurement models. The solution allows for both distance and angle information to be processed algebraically, without iteration, simultaneously

[18]More constraints could be added to the SDP formulation, but this would result in a different solution from the one in [35]. This is corroborate by the results showing a clear gap between the performance of the two algorithms. Similar results stands in the case of minimal angle information, in which it assumed that only the anchor nodes are able to measure one only angle to the target and this set of measurements is then used to create the set of ADOA use in the SDP(AR) algorithm as well as the edge kernel matrix $\mathbf{\Theta}_{\mathrm{E}}$. Indeed even in this case, where the same measured information is given to both algorithms, the E-MDSstill outperforms the SDP approach.

[19]The algorithm is initialized with the centroids used to compute $\hat{\mathbf{\Theta}}_C$.

and at a computational cost of the algorithm is of a few matrix multiplications only, solving a well known bottleneck of the C-MDS solution. We compared the proposed E-MDS algorithm to a constrained optimization technique specifically designed to handle localization in heterogeneous conditions showing the superiority of our solution. More importantly it has been shown that the E-MDS framework allows to exploit the eventual angle information to improve the localization estimate over biased range information. This has been shown to be true even in presence of range only information where the E-MDS still outperforms well established solutions.

APPENDIX A
PROCRUSTES TRANSFORMATION

The procrustes transformation is an *affine transformation*, namely a rigid transformation[20] and a scaling factor .

Given the coordinates of at least $N_A > \eta$ anchor nodes $\mathbf{X}_{N_\mathrm{A}} \in \mathbb{R}^{N_\mathrm{A}\times\eta}$, then the estimated $\hat{\mathbf{X}}$ can be recovered from the output $\mathbf{Y} \in \mathbb{R}^{N\times\eta}$ of a scaling algorithm, *e.g.* the C-MDS, using the following isotropic dilation and the rigid transformations

$$\hat{\mathbf{X}} = \alpha \cdot \mathbf{Y} \cdot \mathbf{F}_\mathrm{R} + \mathbf{c}_\mathrm{D} \otimes \mathbf{1}_N, \tag{46}$$

where the scalar α, the matrix $\mathbf{F}_\mathrm{R} \in \mathbb{R}^{\eta\times\eta}$ and the vector $\mathbf{c}_\mathrm{D} \in \mathbb{R}^{\eta}$ are the coefficients of the Procrustes transformation, calculated as

$$\alpha = \mathrm{tr}\Big([\mathbf{H}_{N_\mathrm{A}}^\mathrm{T} \cdot \mathbf{H}_{N_\mathrm{A}}]^{\frac{1}{2}}\Big) \cdot \frac{\|\bar{\mathbf{X}}_{N_\mathrm{A}}\|_\mathrm{F}}{\|\bar{\mathbf{Y}}_{N_\mathrm{A}}\|_\mathrm{F}}, \tag{47}$$

$$\mathbf{F}_\mathrm{R} = \mathbf{U}_\mathrm{L}^\mathrm{T} \cdot \mathbf{U}_\mathrm{R}, \tag{48}$$

$$\mathbf{c}_\mathrm{D} = \frac{\mathbf{1}_{N_\mathrm{A}}^\mathrm{T}}{N_\mathrm{A}} \cdot \mathbf{X}_{N_\mathrm{A}} - \frac{\alpha \cdot \mathbf{1}_{N_\mathrm{A}}^\mathrm{T}}{N_\mathrm{A}} \cdot [\mathbf{Y}]_{1:N_\mathrm{A},1:\eta} \cdot \mathbf{F}_\mathrm{R}. \tag{49}$$

In the above, $\mathbf{U}_\mathrm{L}$ - $\mathbf{U}_\mathrm{R}$ are the left and right matrices of singular vectors of $\mathbf{H}_{N_\mathrm{A}} = \bar{\mathbf{X}}_{N_\mathrm{A}}^\mathrm{T} \cdot \bar{\mathbf{Y}}_{N_\mathrm{A}} = \mathbf{U}_\mathrm{L} \cdot \mathbf{\Sigma} \cdot \mathbf{U}_\mathrm{R}$, and $\bar{\mathbf{X}}_{N_\mathrm{A}}$ and $\bar{\mathbf{Y}}_{N_\mathrm{A}}$ are the centralized true and rotated coordinate matrices of the anchor nodes, given by

$$\bar{\mathbf{X}}_{N_\mathrm{A}} = \mathbf{X}_{N_\mathrm{A}} - \frac{\mathbf{1}_{N_\mathrm{A}}^\mathrm{T}}{N_\mathrm{A}} \cdot \mathbf{X}_{N_\mathrm{A}} \otimes \mathbf{1}_{N_\mathrm{A}}, \tag{50}$$

$$\bar{\mathbf{Y}}_{N_\mathrm{A}} = \mathbf{Y}_{N_\mathrm{A}} - \frac{\mathbf{1}_{N_\mathrm{A}}^\mathrm{T}}{N_\mathrm{A}} \cdot [\mathbf{Y}]_{1:N_\mathrm{A},1:\eta} \otimes \mathbf{1}_{N_\mathrm{A}}. \tag{51}$$

REFERENCES

[1] G. Wang and H. Chen, "An importance sampling method for tdoa-based source localization," *IEEE Trans. Wireless Commun.*, vol. 10, no. 5, 2011.

[2] B. Liu, H. Chen, Z. Zhong, and H. V. Poor, "Asymmetrical round trip based synchronization-free localization in large-scale underwater sensor networks," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, Nov. 2010.

[3] H. Chen, Q. Shi, R. Tan, H. V. Poor, and K. Sezaki, "Mobile element assisted cooperative localizaitno for wireless sensor networks with obstacles," *IEEE Trans. Wireless Commun.*, vol. 9, no. 3, pp. 956–963, Mar. 2010.

[4] D. Dardari, E. Falletti, and M. Luise, *Satellite and Terrestrial Radio Positioning Techniques - A Signal Processing Perspective*. Academic Press, 2012.

[5] H. Chen, G. Wang, Z. Wang, H. C. So, and H. V. Poor, "Non-line-of-sight node localization based on semi-definite programming in wireless sensor networks," *IEEE Trans. Wireless Commun.*, no. 99, 2012.

[20]Rigid transformations are reflections, rotations and translations [19].

[6] Q. Shi, C. He, H. Chen, and L. Jiang, "Distributed wireless sensor network localization via sequential greedy optimization lagorithms," *IEEE Trans. Signal Process.*, 2010.

[7] Y. Shang and W. Ruml, "Improved MDS-based localization," in *Proc. 2004 IEEE Joint Conference of the Computer and Communications Societies*, vol. 4, pp. 2640–2651.

[8] Z. X. Chen, "A supplement to multidimensional scaling framework for mobile location: a unified view," *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 2030–2034, 2009.

[9] D. Macagnano and G. T. F. de Abreu, "Gershgorin analysis of random Gramian matrices with application to MDS tracking," *IEEE Trans. Signal Process.*, vol. 59, no. 4, pp. 1785–1800, Apr. 2011.

[10] M. Z. Win and R. A. Scholtz, "Impulse radio: how it works," *IEEE Commun. Lett.*, vol. 2, no. 2, pp. 36–38, Feb. 1998.

[11] C. Sun and N. C. Karmakar, "Direction of arrival estimation with a novel single-port smart antenna," *EURASIP J. Applied Signal Process.*, vol. 9, pp. 1364–1375, Sept. 2004.

[12] Y. Shang, W. Ruml, Y. Zhang, and M. P. J. Fromherz, "Localization from mere connectivity," in *2003 International Symposium on Mobile Ad-Hoc Networking and Computing MobiHoc*.

[13] J. C. Gower, "Properties of Euclidean and non-Euclidean distance matrices," *Linear Algebra and its Applications*, vol. 67, pp. 81–97, 1985.

[14] J. Dattorro, *Convex Optimization and Euclidean Distance Geometry*. Meboo Publishing, 2005.

[15] A. Y. Alfakih and H. Wolkowicz, "On the embeddability of weighted graphs in Euclidean spaces," University of Waterloo, Dept. of Combinatorics and Optimization, Tech. Rep., 1998.

[16] K. W. Cheung and H. C. So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 460–470, Feb. 2005.

[17] G. Destino and G. T. F. de Abreu, "Weighing strategy for network localization under scarce ranging information," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3668–3678, July 2009.

[18] C. Fowlkes, S. Belongie, F. Chung, and J. Malik, "Spectral grouping using the Nyström method," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 2, pp. 214–225, Feb. 2004.

[19] T. F. Cox and M. A. A. Cox, *Multidimensional Scaling*, 2nd ed. Chapman & Hall/CRC, 2000.

[20] F.Bavaud, "On the Shoenberg transformations and data analysis: theory and illustrations," arxiv, 2010.

[21] I. J. Schoenberg, "Remarks to Maurice Frechet's article "Sur la definition axiomatique d'une classe d'espace distances vectoriellement applicable sur l'espace de hilbert," *The Annals of Mathematics*, vol. 36, no. 3, pp. 724–732, July 1935.

[22] M. W. Trosset, "Extensions of classical multidimensional scaling: computational theory," *Computational Statistics*, vol. 17, 2002.

[23] K. V. Mardia, "Some properties of classical multi-dimensioinal scaling." *Commun. in Statistics - Theory and Methods*, pp. 1233–1241, 1978.

[24] J. J. Meulman, "The integration of multidimensional scaling and multivariate analysis with optimal transformations," *Psychometrika*, vol. 57, 1992.

[25] A. Y. Alfakih, H. Wolkowicz, and A. Khandani, "Solving Euclidean distance matrix completion problems via semidefinite programming," *J. Computational Optimization and Applications*, vol. 12, no. 1, pp. 13–30, 1999. [Online]. Available: http://portal.acm.org/citation.cfm?id=316254\#

[26] L. Cayton and S. Dasgupta, "Robust Euclidean embedding," in *International Conference on Machine Learning*, I. C. on Machine Learning (ICML), Ed., 2006.

[27] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed. Johns Hopkins University Press, 1996.

[28] G. T. F. de Abreu, "On the simulation of Tikhonov random processes," in *Proc. 2007 IEEE International Conference on Communications*.

[29] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Application to Tracking and Navigation*. Wiley-Interscience, 2001.

[30] G. T. F. de Abreu and G. Destino, "Super MDS: source location from distance and angle information," in *Proc. 2007 Wireless Communications and Networking Conference*, pp. 4430–4434.

[31] O. Besson and L. L. Scharf, "CFAR matched direction detector," *IEEE Trans. Signal Process.*, vol. 54, no. 7, July 2006.

[32] F. Mourad, H. Snoussi, F. Abdallah, and C. Richard, "Anchor-based localization via interval analysis for mobile ad-hoc sensor networks," *IEEE Trans. Signal Process.*, vol. 57, no. 8, pp. 3226–3239, 2009.

[33] S. Venkatesh and R. M. Buehrer, "NLOS mitigation using linear programming in ultrawideband location-aware networks," *IEEE Trans. Veh. Technol.*, vol. 56, no. 5, pp. 3182–3198, 2007.

[34] H. Chen, "Accurate and efficient node localization in wireless sensor networks," Ph.D. dissertation, The University of Tokyo, 2011.

[35] P. Biswas, "Semidefinite programming approaches to distance geometry problems," Ph.D. dissertation, Stanford University, June 2007.

Davide Macagnano (S'06-M'12) obtained his M.Sc. degree from the Politecnico di Milano, Italy in 2005 and the Ph.D. degree from the University of Oulu, Finland in 2012. Since January 2006 he is a researcher at the Centre for Wireless Communications, University of Oulu, Finland. His research interests are Bayesian algorithms, convex optimization and graph theory for statistical signal processing with focus on localization and tracking algorithms. In 2012 he was co-recipient of the Best Paper Award at the Workshop on Positioning Navigation and Communications.

Giuseppe T. F. de Abreu (S'99-M'04-SM'09) is an Associate Professor of Electrical Engineering at Jacobs University, Bremen. He received the B.Eng. degree in Electrical Engineering and a specialization (*Latu Sensu*) degree in Telecommunications Engineering from the Universidade Federal da Bahia (UFBa), Salvador, Bahia, Brazil, in 1996 and 1997, respectively; and the M.Eng. and Ph.D degrees in Physics, Electrical and Computer Engineering from the Yokohama National University, Japan, in March 2001, and March 2004, respectively, having been the recipient of the 2000 Uenohara Award by Tokyo University for his Master's Thesis. Prior to joining Jacobs University he was an Adjunct Professor (Docent) on Statistical Signal Processing and Communications Theory at the Department of Electrical and Information Engineering, University of Oulu, Finland. Dr. Abreu is a member of the Technical Program Committee of various IEEE conferences, an author/co-author of over 130 published technical articles, patents and a few book chapters. He co-authored articles selected as the best paper of the Signal Processing and Communications Track at the Asilomar Conference on Signals Systems and Computers in 2009 and 2012, respectively, and was co-recipient of the Best Paper Award at the Workshop on Positioning Navigation and Communications in 2012. Prof. Abreu is active in various topics ranging from communications theory, signal processing, statistical modeling, localization and tracking algorithms, cognitive radio, wireless security and others.