Multidimensional Scaling Localization with Anchors

Carmelo Di Franco*, Enrico Bini†, Mauro Marinoni*, Giorgio C. Buttazzo*

*ReTiS Laboratory, TeCIP Instutute, Scuola Superiore Sant'Anna, Pisa, Italy

†Department of Computer Science, University of Turin, Italy

Abstract—**Multidimensional Scaling (MDS) is a widely used technique for visualizing a set of objects in an n-dimensional space. It has been extensively applied in wireless sensor networks for deriving the coordinates of a set of nodes in distance-based Localization. Many variants of MDS have been proposed to overcome issues such as partial connectivity and different types of noise in the measurements. In particular, some works adapted and modified the MDS technique to include the notion of anchors. However, in order to maintain the original formulation of MDS, the algorithm was twisted by adding constraints to the minimization function or adapting the final result through roto-translations. Unfortunately, however, these adaptions do not fully solve the problem, because they try to align the relative positions of the nodes to the global reference system provided by the anchors only after the MDS algorithm.**

This paper provides a theoretical generalization of the classical MDS algorithm when some of the coordinates of some elements (e.g., anchors in the case of localization) are known. The proposed generalization can be applied to any of the many MDS variants (e.g., classical MDS, ordinal MDS, MDS-MAP, GM-MDS) that minimize the stress function with the SMACOF technique. The formulation is proved to be correct and does not add any constraints to MDS.

I. INTRODUCTION

One of the main issues in wireless sensor networks is to locate the positions of the nodes. In indoor systems, where the GPS is not available, a common solution is the distance-based localization, which derives the coordinates by measuring the inter-nodes distances. There exist different techniques that estimate the node coordinates from distance measurements, such as trilateration, multilateration, and Multidimensional Scaling (MDS). MDS [1], [2] is a technique that aims at visualizing a set of objects in an n-dimensional space. It takes as input a Dissimilarity Matrix that expresses how much two objects are dissimilar along one quality and finds a set of coordinates such that the distance between each couple of objects is proportional to the value of dissimilarity. MDS has been used for plotting sets of data in many application fields, such as economics and psychology. In the last decades, MDS has also been used for localization, where the objects are the nodes, and the dissimilarity matrix contains the inter-node distances.

In wireless sensor networks, the use of anchor nodes deeply impacts the accuracy of the entire localization system. However, in the MDS formulation the notion of anchor nodes is missing, since such a technique was designed

for plotting generic objects with qualitative characteristics. Moreover, some limitation such as partial connectivity [3], [4] and different type of noise exist in the specific case of network localization [5]. For this reason, many variants of MDS have been proposed in the literature. Referring to the specific problems of MDS with anchors, some works proposed solutions for including the known coordinates into the algorithm. Such variants of MDS can be distinguished in two categories: distributed and centralized approaches. An approach is considered distributed when the computation of the algorithm is divided among the components of the network.

A distributed version of MDS, called distributed weighted Multidimensional Scaling (dwMDS), was proposed by Costa et. al [6]. They successfully introduced the notion of anchors and were able to consider them in the MDS formulation since they split the computation on each node without modifying the anchor's coordinates.

When referring to centralized approaches, a common solution is to apply a roto-translation transformation after the MDS computation and superimpose the estimated coordinates over the anchors [3], [4], [7], [8]. Instead, Biaz and Ji [9] used a different approach consisting of updating the anchor's positions during the minimization procedure. However, contrarily to the current MDS versions, the positions of the nodes whose coordinates are known (anchor nodes) should not be modified during the minimization. For all these cases, at the end of the minimization, the coordinates of the anchors slightly changes due to the minimization procedure, leading to a position error that reduces the overall accuracy of the system. Also, in the case where the anchor coordinates are modified at the end of the minimization, the output is not exact since the anchor's coordinates are not fully used for finding the best estimation.

On the other hand, as previously stated, dwMDS provides a valid solution to the problem. However, a distributed approach is often used in very large sensors networks usually composed of inexpensive sensors with low computation capabilities. Moreover, such sensors are typically static. In this case, the dwMDS distributes the computation among all the sensors with the only drawback of a high exchange of wireless communication (and hence, longer time needed) in order to converge to the final coordinates. Centralized approaches, instead, are usually preferred in small networks and in applications which consider node mobility, e.g., in the case of a small team of robots [10]. Another application case in which a centralized approach is commonly used includes indoor people tracking

[11].

This work provides a theoretical generalization of the classical MDS algorithm when the coordinates of some variables (e.g., anchors) are known. We propose a modified MDS formulation to include the notion of anchors directly. Our generalization can be used to any of the many MDS variants (e.g., MDS, wMDS, ordinal MDS, MDS-MAP, GM-MDS). Our formulation does not add any constraints to MDS nor requires additional computation, such as roto-translations or tailored procedures.

II. RELATED WORK

Several authors attempted to extend MDS localization for including the knowledge of anchors. However, most of such extensions suffer from two drawbacks: the MDS minimization does not benefit from anchors information, and the algorithm introduces noise on the coordinates.

Two different approaches have been used to fix the problem described above: a) apply a roto-translation to the MDS coordinates, b) act on the minimization process by adding constraints or modifying the cost function. This section explains why both approaches do not solve the problem efficiently. In this work, we refer only to centralized approaches, distributed solution has been proposed by Costa et. al. [6].

A. Applying a roto-translation transformation

The first class of works divides the localization algorithm in two step: first, the relative coordinates are estimated through MDS; then, a roto-translation transformation is applied at the end of the minimization to align the estimated coordinates of the anchors to their real positions.

For example, Ji and Zha [7] align the relative positions of the nodes to their real positions by computing a shift, rotation, and reflection of the coordinates. They compute the rotation and translation matrices by the aid of at least three nodes. Similarly, in the MDS-MAP algorithm [3], the authors transform the global map to an absolute map based on the absolute positions of the anchors. They state that for m anchors, the complexity of this step is $O(m^3 + n)$.

Latsoudas and Sidiropoulos [12] avoid the problem of aligning the relative positions of the nodes to the real location of the anchors by carefully placing the anchors to form an orthogonal triangle, and the orthogonal sides of this triangle are chosen as coordinate basis vectors. Then, all the projections are computed directly onto the native coordinate basis, thus, avoiding the need of applying a roto-translation. This approach has the drawback that the anchor nodes need to be placed in the environment according to a specific pattern.

Cheung and So [13] proposed an approach that uses the anchor coordinates to compute the rotation matrix. Their approach is initially similar to the one proposed in this paper. However, we do not compute any rotation matrix. Moreover, their MDS formulation is specific to the case in which only one node has to be located through the use of m anchors.

B. Modifying the MDS minimization or adding constraints

Biaz and Ji [9] described a way of updating the anchor's positions during the minimization procedure. They first run the MDS step and then, after a check of communication constraints, they update the anchors' position. Note that this procedure is a workaround since all nodes are treated as unknown, and the MDS minimization is also performed on the anchors. After the minimization, the position of the anchors is updated, but it is also necessary to adjust the distances between the anchors and the nodes to maintain consistency.

Another technique that exists in the literature of MDS consists of applying generic restrictions to the variables [14]. In particular, information regarding the anchors locations can be included as a constraint of MDS. However, such an approach increases the space of the variables and leads to higher complexity and computational cost. Conversely, reducing the space of the variables to only the unknown nodes and consider the anchors as a constant value, not only improves the precision but also reduces the complexity of the algorithm. This is the intuition behind our approach. With respect to the related work, our formulation provides a mathematical formulation to estimate only the coordinates of the nodes. Since the anchor's coordinates are used to estimate the node positions, the two presented drawbacks do not subsist in our approach, leading to a higher accuracy in the localization.

III. REVIEW OF MULTIDIMENSIONAL SCALING THEORY

Multidimensional Scaling (MDS) is a technique that represents a set of elements in an n-dimensional space using the similarities/dissimilarities between pairs of elements as distance information. There exist several variants of MDS such as Classical MDS, Metric MDS, Non-Metric MDS, depending on the characteristics of the distance information. Given a network composed by N nodes in a d-dimensional space, whose coordinates $\mathbf{X} = [\mathbf{x}_1, \cdots, \mathbf{x}_N]^T \in \mathbb{R}^{N \times d}$ are unknown, the algorithm recovers the coordinates of the elements by minimizing the mismatch of the following function

$$\min_{\mathbf{X}} \mathcal{S}(\mathbf{X}) = \min_{X} \sum_{i < j \leq N} w_{ij} \left(\hat{d}_{ij} - d_{ij}(\mathbf{X}) \right)^2 \quad (1)$$

where w_{ij} is a weight defining the quality of the measurement $\hat{d}_{ij}$ and $d_{ij}(\mathbf{X}) = \|\mathbf{x}_i - \mathbf{x}_j\|$ is the Euclidean distance.

The objective function $\mathcal{S}$, also called stress-function, can be minimized in different ways, such as using the steepest descent approach. De Leeuw [15] proposed an iterative method that at each step minimizes a simple convex function which majorizes the complex function. This approach is called "Scaling by MAjorizing a COmplicated Function" (SMACOF) and it was proved to perform significantly better with respect to other approaches [15], in terms of guarantees and rate of convergence.

The non linear least squares problem in Equation (1) is solved minimizing iteratively a convex function $\mathcal{T}(\mathbf{X}, \mathbf{Z}) \geq \mathcal{S}(\mathbf{X})$. $\mathcal{T}$ bounds $\mathcal{S}$ from above and touches the surface of $\mathcal{S}$ at point $\mathbf{Z}$. The iterative procedure is summarized in Algorithm 1.

Algorithm 1 Scaling by MAjorizing a COmplicated Function (SMACOF)

---

input: initial position estimate $\mathbf{X}^{(0)}$
        stress function $\mathcal{S}$
        majorizing function $\mathcal{T}$
repeat
        $\mathbf{Z} = \mathbf{X}^{k-1}$
        $\mathbf{X}^k \leftarrow \min_{\mathbf{X}} \mathcal{T}(\mathbf{X}, \mathbf{Z})$
until $\mathcal{S}(\mathbf{X}^{k-1}) - \mathcal{S}(\mathbf{X}^k) < \epsilon$

A. The majorization function

The stress function $\mathcal{S}$ can be expanded as follows:

$$\mathcal{S}(\mathbf{X}) = \sum_{i<j\leq N} w_{ij}\left(\hat{d}_{ij} - d_{ij}(\mathbf{X})\right)^2 =$$
$$= \sum_{i<j} w_{ij}\hat{d}_{ij}^2 + \sum_{i<j} w_{ij}d_{ij}^2(\mathbf{X}) - 2\sum_{i<j} w_{ij}\hat{d}_{ij}d_{ij}(\mathbf{X}).$$

Note that the first term is a constant and the second term is quadratic in $\mathbf{X}$ and therefore relatively easily solved. The third term is bounded using the Cauchy-Schwarz inequality using the fact that:

$$d_{ij}(\mathbf{X}) = \|\mathbf{x}_i - \mathbf{x}_j\| = \|\mathbf{x}_i - \mathbf{x}_j\| \frac{\|\mathbf{z}_i - \mathbf{z}_j\|}{\|\mathbf{z}_i - \mathbf{z}_j\|} \geq$$
$$\frac{(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{z}_i - \mathbf{z}_j)}{\|\mathbf{z}_i - \mathbf{z}_j\|} \quad (2)$$

where $\mathbf{Z} = [\mathbf{z}_1, \cdots, \mathbf{z}_N]^T \in \mathbb{R}^{N \times d}$. Hence, the third term can be bounded by:

$$\sum_{i<j} w_{ij}\hat{d}_{ij}d_{ij}(\mathbf{X}) \geq \sum_{i<j} w_{ij}\hat{d}_{ij}\frac{(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{z}_i - \mathbf{z}_j)}{||\mathbf{z}_i - \mathbf{z}_j||} \quad (3)$$

Thus, we have a simple quadratic function $\mathcal{T}(\mathbf{X}, \mathbf{Z})$ that majorizes the stress:

$$\mathcal{S}(\mathbf{X}) \leq \mathcal{T}(\mathbf{X}, \mathbf{Z}) = \sum_{i<j} w_{ij}\hat{d}_{ij}^2 + \sum_{i<j} w_{ij}d_{ij}^2(\mathbf{X})$$
$$- 2\sum_{i<j} w_{ij}\hat{d}_{ij}^2\frac{(\mathbf{x}_i - \mathbf{x}_j)^T(\mathbf{z}_i - \mathbf{z}_j)}{||\mathbf{z}_i - \mathbf{z}_j||} \quad (4)$$

$\mathcal{T}(\mathbf{X}, \mathbf{Z})$ can be written in matrix form:

$$\mathcal{T}(\mathbf{X}, \mathbf{Z}) = \mathbf{C} + \operatorname{tr}(\mathbf{X}^T \mathbf{V} \mathbf{X}) - 2\operatorname{tr}(\mathbf{X}^T \mathbf{B}(\mathbf{Z})\mathbf{Z}) \quad (5)$$

where $\mathbf{V}$ and $B(\mathbf{Z})$ are two matrices whose elements are defined as follow:

$$v_{ij} = \begin{cases} \displaystyle\sum_{k=1,k\neq j}^{N} w_{kj} & \text{if } i \neq j, \\ \displaystyle\sum_{k=1,k\neq j}^{N} v_{kj} & \text{if } i = j. \end{cases}$$

$$b_{ij} = \begin{cases} \displaystyle\sum_{k=1,k\neq j}^{N} w_{kj}\frac{\hat{d}_{ij}}{d_{ij}(\mathbf{Z})} & \text{if } i \neq j, \\ \displaystyle\sum_{k=1,k\neq j}^{N} b_{kj} & \text{if } i = j. \end{cases}$$

Thus, the minimum of the function can be computed as:

$$\mathbf{X} = \min_{\mathbf{X}} \mathcal{T}(\mathbf{X}, \mathbf{Z}) = \mathbf{V}^{-1}\mathbf{B}(\mathbf{Z})\mathbf{Z} \quad (6)$$

IV. PROPOSED APPROACH

Let us consider a set of $N = n + m$ nodes where the first n nodes have unknown positions and the last m are the anchor nodes. Referring to the matrices in Equation (6), we can partition $\mathbf{X}$ and $\mathbf{Z}$ as follows:

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_u \\ \mathbf{X}_a \end{bmatrix}, \quad \mathbf{Z} = \begin{bmatrix} \mathbf{Z}_u \\ \mathbf{Z}_a \end{bmatrix}, \quad \text{with}$$
$$\mathbf{X}_u = [\mathbf{x}_1, \cdots, \mathbf{x}_n]^T \in \mathbb{R}^{n \times d}$$
$$\mathbf{X}_a = [\mathbf{x}_{n+1}, \cdots, \mathbf{x}_{n+m}]^T \in \mathbb{R}^{m \times d}$$
$$\mathbf{Z}_u = [\mathbf{z}_1, \cdots, \mathbf{z}_n]^T \in \mathbb{R}^{n \times d}$$
$$\mathbf{Z}_a = [\mathbf{z}_{n+1}, \cdots, \mathbf{z}_{n+m}]^T \in \mathbb{R}^{m \times d}.$$

Similarly, we can partition $\mathbf{V}$ and $\mathbf{B}(\mathbf{Z})$ in blocks as follow:

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{12}^T & \mathbf{V}_{22} \end{bmatrix}, \mathbf{B}(\mathbf{Z}) = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}^T & \mathbf{B}_{22} \end{bmatrix}, \quad (7)$$

where matrices $\mathbf{V}_{11}, \mathbf{B}_{11}$ are of size $n \times n$, $\mathbf{V}_{12}, \mathbf{B}_{12}$ are $m \times n$, and $\mathbf{V}_{22}, \mathbf{B}_{22}$ are $m \times m$.

The following theorem provides a way for computing $\mathbf{X}_u$ as a function of the anchors coordinates $\mathbf{X}_a$.

Theorem 1. *Given the stress function $\mathcal{T}(\mathbf{X}, \mathbf{Z})$ of Equation (5), if we know the exact value of the subset $\mathbf{X}_a \in \mathbf{R}^{d \times m}$ of $\mathbf{X}$, it is possible to compute the remaining unknown values $\mathbf{X}_u$ as a function of $\mathbf{X}_a$:*

$$\mathbf{X}_u = \mathbf{V}_{11}^{-1}\left(\mathbf{B}_{11}\mathbf{Z}_u + \mathbf{B}_{12}\mathbf{Z}_a - \mathbf{V}_{12}\mathbf{X}_a\right), \quad (8)$$

with $\mathbf{V}_{11}$, $\mathbf{V}_{12}$, $\mathbf{B}_{11}$, and $\mathbf{B}_{12}$ matrix block defined in (7).

Proof. See Appendix A.   □

The major practical benefit deriving from this theorem is the possibility to extend any variant of MDS that uses the SMACOF implementation to support the notion of anchors by simply modifying the computation of $\mathbf{X}$.

Note that, Equation (8) reduces to Equation (6) in the absence of anchors ($m = 0$). Hence, our approach is not limited to the a-priori knowledge on the anchors, but can straightforwardly be used also when their number varies.

V. EXPERIMENTAL EVALUATION

The proposed formulation, named anchored MDS, has been evaluated in comparison with a generic MDS approach that applies a roto-translation transformation to the coordinates estimated by the minimization. It is worth noting that the techniques presented in the related work mainly differ in the way they compute the roto-translation matrix. Different techniques lead to different accuracy. However, despite the specific algorithm used to align the coordinates, the two drawbacks highlighted in Section II are common to all those approaches. We simulated a network with a varying number of unknown nodes and anchors. The noise in the distance measurement between two nodes has been modeled as:

$$\hat{d}_{ij} = d_{ij}(X) + \nu$$

where $\nu \sim \mathcal{N}(\mu, \sigma^2)$ is a Gaussian noise with mean μ and standard deviation σ.

In all the simulations, the network is fully connected. The problem of partial connectivity has been addressed in detail in the literature (examples are [4], [8], [3]). However, since we are evaluating a generalization of the classical MDS, that can also be applied to all these variations, we are not considering the case of partial connectivity and focus only on the case of full connectivity in which we know the positions of some anchors. Moreover, since the value of the weights w_{ij} associated to each distance measurements affects the minimization, we decided to set all the weights to 1.

To evaluate the goodness of the approaches with respect to the real coordinates we use as a metric the Root Mean Square Error (RMSE) between true and estimated node positions, defined as:

$$\text{RMSE}\left(\mathbf{X}, \hat{\mathbf{X}}\right) := \sqrt{\frac{1}{n} \sum_{i=1}^{n} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2}. \qquad (9)$$

A. A graphic example

Figure 1 and 2 show a trivial example in which a network with $m = 3$ anchors and $n = 2$ nodes are localized with the classical MDS and the anchored MDS proposed in this paper. The noise on the distance measurements ν has a zero mean $\mu = 0$ and a standard deviation $\sigma = 1$. Figure 1 shows the real and estimated coordinates computed with the classical MDS. At the end of the minimization, a roto-translation is applied in order to match the anchors locations with the estimated ones. The red dots represent the estimated coordinates during the SMACOF iterations. Note that, for visualization purposes, the roto-translation has also been applied to the coordinates computed during the SMACOF iterations. In Figure 2, the coordinates are estimated with our formulation. As it can be seen from the figure, the anchor nodes are fixed at the correct position from the beginning and only the coordinates of the n nodes are computed during the SMACOF iterations.

The two drawbacks described in Section II are clearly highlighted through the figures. First, the classical MDS formulation does not benefit from the anchor information. The
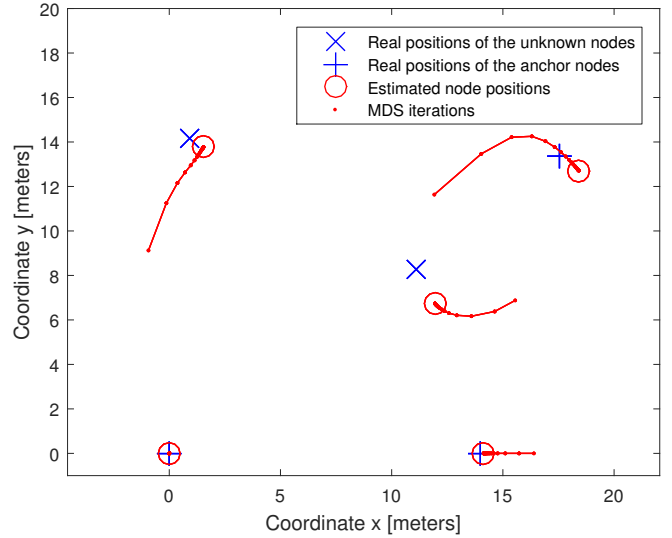

Fig. 1. Estimated coordinates computed with the classical MDS. The red dots represent the coordinates during the SMACOF iteration.
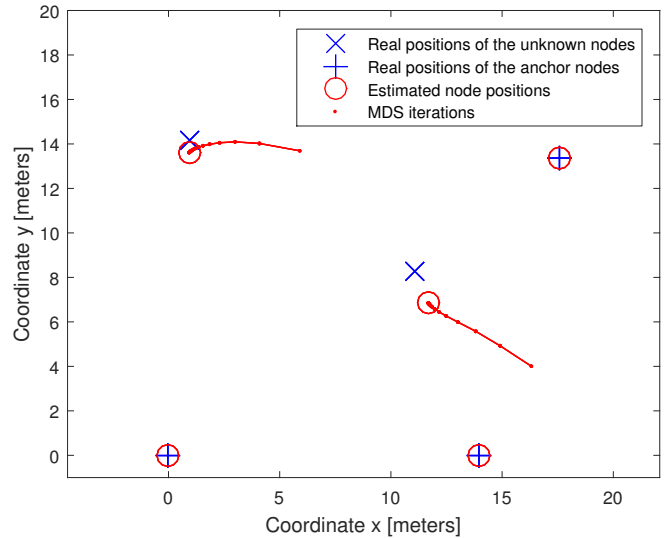

Fig. 2. Estimated coordinates computed with our formulation MDS with anchors. Only the unknown node coordinates are estimated leading to a reduction of the overall error.

distances between anchors are set to the exact values since such an information is known. However, their coordinates are subject to a minimization even if it is not required. Moreover, the knowledge of the exact position of the anchors cannot be properly used during the MDS minimization.

Second, the algorithm introduces noise on the coordinates. As it can be seen from Figure 1, the anchors location is not precisely reconstructed, even if such an information is known.

B. Experimental results

Some experiments have also been carried out to show how the overall accuracy varies as a function of the error in the distance measurements. In the first experiment, we verified
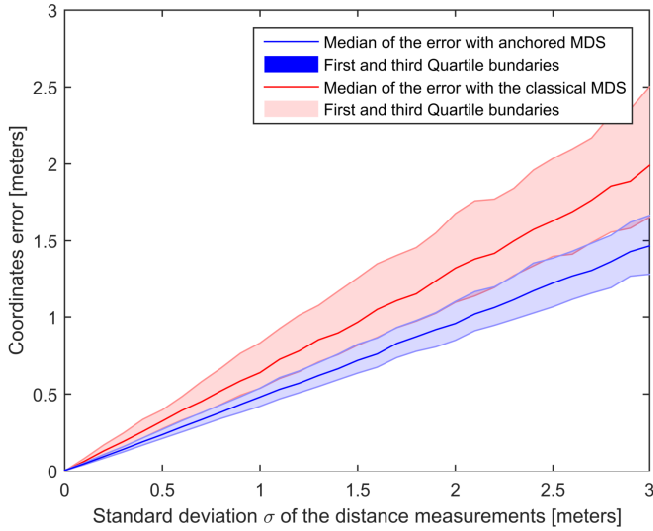
Fig. 3. Comparison between the error in the coordinates (RMSE) computed with anchored MDS and the classical MDS approach + roto-translation as a function of the standard deviation σ.

Fig. 4. Comparison between the error in the coordinates (RMSE) computed with anchored MDS and the classical MDS approach + roto-translation varying the Anchors/Nodes ration.

how the overall accuracy varies as a function of the error in the distance measurements. We varied σ from 0 to 3 meters, with a step of 0.1 meters. For each vale of σ, we run 1000 simulations in which we chose **X** randomly. The network is composed of $N = 50$ nodes of which $m = 5$ anchors. Figure 3 shows the median of the RMSE error (bold lines), while the colored surfaces are delineated by the first and third quartile of the RMSE error. As it can be seen from the figure, our approach always outperforms the classical MDS approach. This was expected since the anchors locations in our approach are placed at the exact position and such an information positively affects the other nodes location during the SMACOF iterations, leading to a lower RMSE.

The second experiment shows how the number of anchors affects the accuracy of our approach. We simulated a network of $N = 50$ nodes, where the distance measurements had a standard deviation $\sigma = 1.5$ and varied the number of anchors m from 3 to N. For each m we run 1000 iterations.

As expected, both approaches decrease monotonically to zero as the number of anchors increases. However, since our formulation takes advantage of the anchor's location during the minimization, it outperforms the classical approach achieving both a smaller mean error and variance.

Finally, we performed a third experiment in which we varied the number of nodes N while keeping a fixed ratio $m/N = 20\%$. As in the second experiment, the standard deviation of the distance measurement is fixed to $\sigma = 1.5$, and for each value of N we run 1000 iterations.

As shown in Figure 5, both approaches present a monotonic decrease with the number of nodes, and the error of the anchored MDS is always below the classical MDS. This is an expected consequence of the fully connected network since the number of links increases proportionally with nodes.
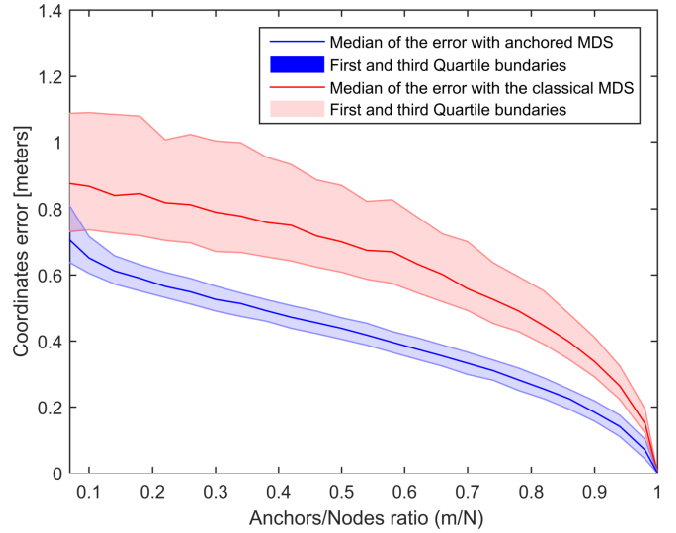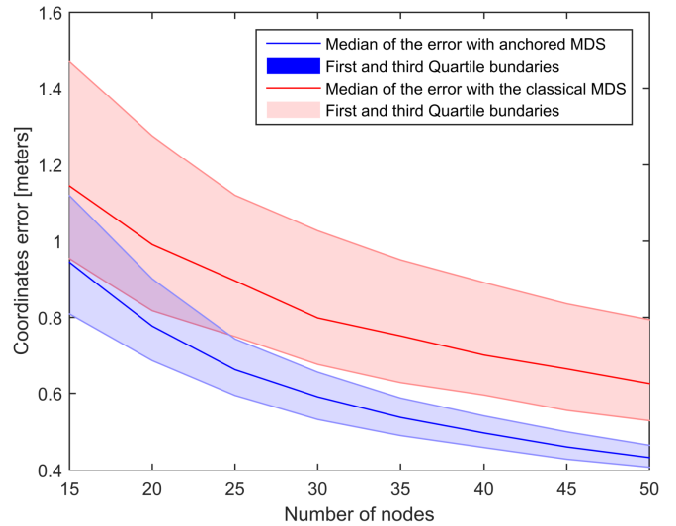

Fig. 5. Comparison between the error in the coordinates (RMSE) computed with anchored MDS and the classical MDS approach + roto-translation varying the number of nodes and fixing the Anchors/Nodes ratio at 20%.

VI. DISCUSSION

A common approach used to align relative and global coordinates is the Procrustes analysis [2], which is a method used to analyze two sets of elements. It finds a linear transformation (translation, reflection, orthogonal rotation, and scaling) of the points in matrix B to best conform them to the points in matrix A. In particular, the Orthogonal Procrustes does not scale the matrix B. The orthogonal matrix R can be computed as:

$$R = \arg\min_{\Omega} \|\Omega A - B\|_F \quad \text{subject to} \quad \Omega^T\Omega = I \quad (10)$$

where $\|\cdot\|_F$ denotes the Frobenius norm. Despite the overall error computed with such an approach leads to the minimum

error, the coordinates of the anchor will never be aligned to their real positions. Moreover, in the case of node mobility, it will be impossible to correlate the node coordinates at successive MDS computation, since the rotation between the nodes will differ at every step. This will report erroneous movements of the nodes unrelated with their actual dynamic. Thus, the Procrustes analysis should be used only in the case of relative localization and not when anchors information is available.

VII. CONCLUSIONS

In this paper, we proposed a theoretical generalization of the Classical MDS, and all its variants, introducing the notion of anchors directly in the SMACOF formulation. The proposed formulation correctly minimizes the stress function by reducing the number of variables, while other approaches force the output of the classical MDS to be consistent to the anchors coordinates. Unfortunately, however, these adaptations do not fully solve the problem, because they try to align the relative positions of the nodes to the global reference system proved by the anchors only after the MDS algorithm. Thanks to its simplicity and straightforward applicability, the proposed approach can effectively be used to modify any MDS variant to include the concept of anchors with minor modifications to the SMACOF implementation.

REFERENCES

[1] J. B. Kruskal, "Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis," *Psychometrika*, vol. 29, no. 1, pp. 1–27, 1964.

[2] I. Borg and P. Groenen, *Modern Multidimensional Scaling: Theory and Applications*, ser. Springer Series in Statistics. Springer, 2005.

[3] Y. Shang and W. Ruml, "Improved mds-based localization," in *INFO-COM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4. IEEE, 2004, pp. 2640–2651.

[4] Y. Shang, W. Rumi, Y. Zhang, and M. Fromherz, "Localization from connectivity in sensor networks," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 15, no. 11, pp. 961–974, 2004.

[5] C. Di Franco, A. Prorok, N. Atanasov, B. Kempke, P. Dutta, V. Kumar, and G. J. Pappas, "Uwb network localization with non-line-of-sight range measurements," in *Proceedings of the 16th International Conference on Information Processing in Sensor Networks*. IEEE Press, 2017, p. accepted for publication.

[6] J. A. Costa, N. Patwari, and A. O. Hero III, "Distributed weighted-multidimensional scaling for node localization in sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 2, no. 1, pp. 39–64, 2006.

[7] X. Ji and H. Zha, "Sensor positioning in wireless ad-hoc sensor networks using multidimensional scaling," in *INFOCOM 2004. Twenty-third AnnualJoint Conference of the IEEE Computer and Communications Societies*, vol. 4. IEEE, 2004, pp. 2652–2661.

[8] A. Amar, Y. Wang, and G. Leus, "Extending the classical multidimensional scaling algorithm given partial pairwise distance measurements," *Signal Processing Letters, IEEE*, vol. 17, no. 5, pp. 473–476, 2010.

[9] S. Biaz and Y. Ji, "Precise distributed localization algorithms for wireless networks," in *Sixth IEEE international symposium on a world of wireless mobile and multimedia networks*. IEEE, 2005, pp. 388–394.

[10] L. Oliveira, H. Li, L. Almeida, and T. E. Abrudan, "Rssi-based relative localisation for mobile robots," *Ad Hoc Networks*, vol. 13, pp. 321–335, 2014.

[11] F. Cabero, Jose Maria sand De la Torre, A. Sanchez, and I. Arizaga, "Indoor people tracking based on dynamic weighted multidimensional scaling," in *Proceedings of the 10th ACM Symposium on Modeling, analysis, and simulation of wireless and mobile systems*, 2007, pp. 328–335.

[12] G. Latsoudas and N. D. Sidiropoulos, "A fast and effective multidimensional scaling approach for node localization in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 5121–5127, 2007.

[13] K. W. Cheung and H.-C. So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE transactions on signal processing*, vol. 53, no. 2, pp. 460–470, 2005.

[14] J. De Leeuw and W. J. Heiser, "Multidimensional scaling with restrictions on the configuration," *Multivariate analysis*, vol. 5, pp. 501–522, 1980.

[15] J. De Leeuw, "Applications of convex analysis to multidimensional scaling," *Department of Statistics, UCLA*, 2005.

APPENDIX A
PROOF OF THEOREM 1

The trace of a matrix is defined as the sum of the elements along the diagonal. Hence, the trace of the $d \times d$ matrix $\mathbf{X}^T \mathbf{V} \mathbf{X}$ of (5) is equal to

$$\operatorname{tr}(\mathbf{X}^T \mathbf{V} \mathbf{X}) = \sum_{k=1}^{d} \mathbf{X}_{(k)}^T \mathbf{V} \mathbf{X}_{(k)}$$

with $\mathbf{X}_{(k)} \in \mathbb{R}^{N \times 1}$ being the k-th column vector of $\mathbf{X}$. Hence, we can rewrite Equation (5) as follows:

$$\mathcal{T}(\mathbf{X}, \mathbf{Z}) = \mathbf{C} + \operatorname{tr}(\mathbf{X}^T \mathbf{V} \mathbf{X}) - 2 \operatorname{tr}(\mathbf{X}^T \mathbf{B}(\mathbf{Z}) \mathbf{Z})$$
$$= \mathbf{C} + \sum_{k=1}^{d} \mathbf{X}_{(k)}^T \mathbf{V} \mathbf{X}_{(k)} - 2 \sum_{k=1}^{d} \mathbf{X}_{(k)}^T \mathbf{B}(\mathbf{Z}) \mathbf{Z}_{(k)}.$$

By partitioning in block the above expression, we find

$$\mathcal{T}(\mathbf{X}, \mathbf{Z}) = \mathbf{C} + \sum_{k=1}^{d} \mathbf{X}_{(k)}^T \mathbf{V} \mathbf{X}_{(k)} - 2 \sum_{k=1}^{d} \mathbf{X}_{(k)}^T \mathbf{B}(\mathbf{Z}) \mathbf{Z}_{(k)}$$
$$= \mathbf{C} + \sum_{k=1}^{d} \left(\mathbf{X}_{u(k)}^T \mathbf{V}_{11} \mathbf{X}_{u(k)} + 2 \mathbf{X}_{u(k)}^T \mathbf{V}_{12} \mathbf{X}_{a(k)} \right.$$
$$\left. + \mathbf{X}_{a(k)}^T \mathbf{V}_{22} \mathbf{X}_{a(k)} \right)$$
$$- 2 \sum_{k=1}^{d} \left(\mathbf{X}_{u(k)}^T \mathbf{B}_{11} \mathbf{Z}_{u(k)} + 2 \mathbf{X}_{u(k)}^T \mathbf{B}_{12} \mathbf{Z}_{a(k)} \right.$$
$$\left. + \mathbf{X}_{a(k)}^T \mathbf{B}_{22} \mathbf{Z}_{a(k)} \right). \tag{11}$$

By differentiating the expression of $\mathcal{T}(\mathbf{X}, \mathbf{Z})$ of Equation (11) with respect to the unknowns $\mathbf{X}_{u(k)}$ only, we get:

$$\frac{\partial \mathcal{T}(\mathbf{X}, \mathbf{Z})}{\partial \mathbf{X}_{u(k)}} = 2(\mathbf{V}_{11} \mathbf{X}_{u(k)} + \mathbf{V}_{12} \mathbf{X}_{a(k)}$$
$$- \mathbf{B}_{11} \mathbf{Z}_{u(k)} - \mathbf{B}_{12} \mathbf{Z}_{a(k)}), \quad (12)$$

and by setting it equal to $\mathbf{0}$ we can find the unknowns $\mathbf{X}_{u(k)}$, which are

$$\mathbf{X}_{u(k)} = \mathbf{V}_{11}^{-1} \left(\mathbf{B}_{11} \mathbf{Z}_{u(k)} + \mathbf{B}_{12} \mathbf{Z}_{a(k)} - \mathbf{V}_{12} \mathbf{X}_{a(k)} \right),$$

or, equivalently, in a matrix form

$$\mathbf{X}_u = \mathbf{V}_{11}^{-1} \left(\mathbf{B}_{11} \mathbf{Z}_u + \mathbf{B}_{12} \mathbf{Z}_a - \mathbf{V}_{12} \mathbf{X}_a \right). \tag{13}$$