20

Data-Processing and Optimization Methods for Localization-Tracking Systems

Giuseppe Destino, Davide Macagnano and Giuseppe Abreu
University of Oulu - Centre for Wireless Communications
Finland

1. Introduction

During the last years the increasing demand for location-based services (LBS) stimulated the developments of localization-tracking (LT) technologies. In this regard, a widely known and utilized LT technology is given by the Global Positioning System (GPS), that, based on satellite communications can achieve accuracy of tens of centimeters. However the inability to receive the satellite signal in indoor environments strongly limits the usage of GPS technology to outdoor scenarios. As a consequence, future indoor positioning services, *e.g.* surveillance, logistics and remote health-care applications, need to rely on the development of new LT systems based on short/medium-range wireless technologies. In this regard, thanks to their ability to provide accurate distance measurements under both LOS and non-LOS channel conditions, ultra-wideband (UWB) and chirp spread spectrum (CSS) technologies are two promising technologies to enable indoor LT services. In this chapter, we focus on technological and theoretical aspects of LT systems for indoor applications. In particular we consider two fundamental functionalities, namely, the data processing and the LT algorithm. We start describing an efficient wavelet-based filtering technique to process the data and to assess their reliability and we show how estimates on the confidence on the measurements can be used to improve the target locations. In the second part of the chapter we first cover some well known non-parametric state-of-the-art solutions to the localization problem, namely the classical-multidimensional scaling (MDS), the Nyström approximation and the SMACOF algorithm. Following we propose a novel low-complex technique that goes under the name of linear global distance continuation (L-GDC) and that we show to achieve the same performance of the maximum-likelihood estimator. Finally, the chapter ends with the simulation results and a short discussion on the open challenges.

2. Overview of a location-tracking system

In the near future it is expected that LT services will find usage in a very diversified range of scenarios. As an example, in the office environment illustrated in figure 1 it can be imagined that a LT system will be able to localize mobile or static terminals, *e.g.* notebooks, printers, PDAs and smart-phones using the radio access provided by an existing wireless network infrastructure (*e.g.* UWB,Wi-Fi, LTE-A, and so forth).
Assuming that all devices can communicate, up to their maximum radio range, amongst themselves and that a server is available to run the LT engine, the LT problem reduces to

Source: Communications and Networking, Book edited by: Jun Peng,
ISBN 978-953-307-114-5, pp. 434, September 2010, Sciyo, Croatia, downloaded from SCIYO.COM

find the position of the nodes in the network given a subset of the possible measurements amongst the devices.
To design a LT system, the first distinction that needs to be made is between the anchor nodes, whose location is assumed to be known to the system, and the target nodes, whose position needs to be estimated. For instance, referring to the office scenario depicted in figure 1, the N_A radio access points can be associated to the anchor nodes and the remaining N_T terminal devices with the targets.
At this point it is possible to introduce the three essential functional blocks characterizing of a standard LT system, namely, data acquisition, data processing and localization engine block illustrated in figure 2. The data acquisition block deals with the problem of extracting physical parameters such time-of-arrival (ToA), time-difference-of-arrival (TDoA), received-signal-strenght (RSS) or angle-of-arrival (AoA) Mao et al. (2007) from the radios.
As shown in Li & Pahlavan (2004); Joon-Yong & Scholtz (2002), time-based metrics can provide distance measurements with few centimeters of error in line-of-sight (LOS) and tens of centimeters in non-LOS (NLOS) channel conditions.
Regarding AoA measurements, they are typically based on beam-forming or sensor array signal processing methods and they can achieve accuracies of few degrees in LOS conditions. In contrast to time-based mechanisms, AoA-enabled technology employs narrowband signal, and it is used to develop LT systems in large and open space environments.

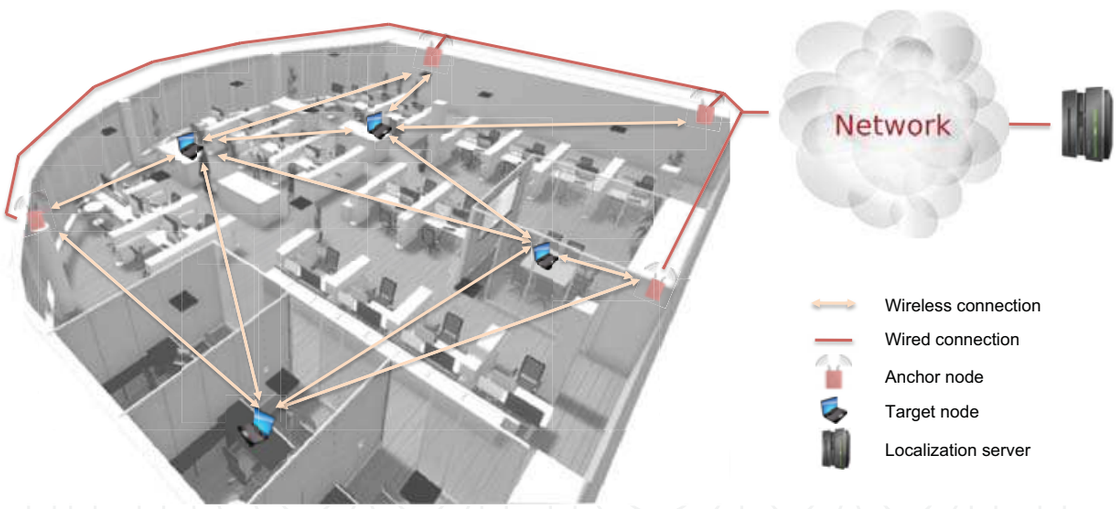


Fig. 1. Typical LT system in indoors

The third type of information is the RSS, which is a power-based metric and typically, available for both narrowband and wideband radio techologies. However the major problems with RSS measurements is their sensitivity to fading and the channel propagation model used. For these reasons, power-based LT systems are inaccurate, especially if RSS's are used to measure distances. Table 1 summarizes the features of the aforementioned metrics.
The second block in figure 2 deals with the processing of the observed data. In this block, a time-series filter is typically conceived to improve the signal to noise ratio of the measurements. The choice of the technique, however, depends on the dynamic and noise models assumed in the scenarios.

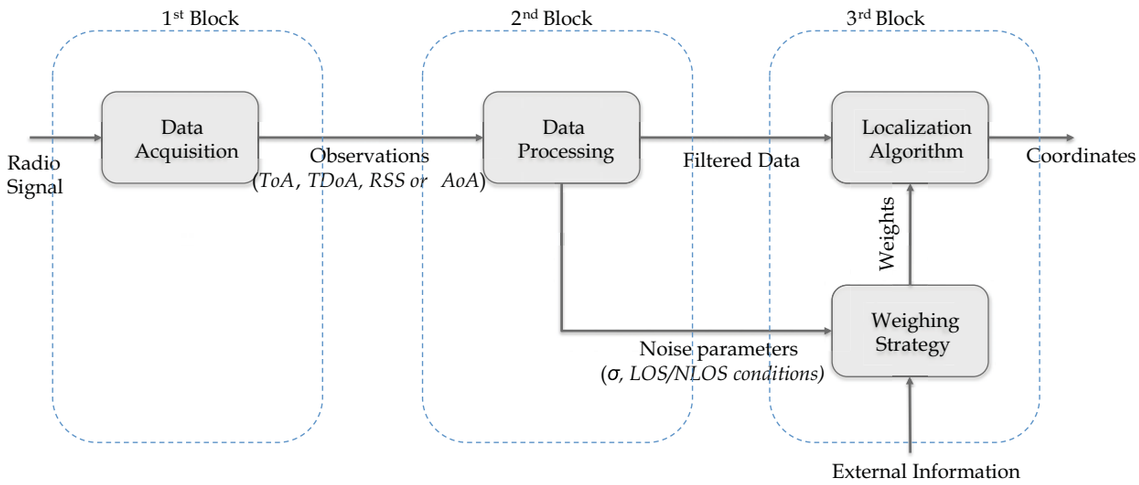


Fig. 2. Functional blocks of an LT system

For instance, in the case of a quasi-static target nodes and a stationary zero mean Gaussian noise affecting the observations even a simple moving average filter can substantially improve the location accuracy of the LT system.

Information Type	Strength	Weakness	Technology
ToA	High precision High multi-path resolution	Synchronization	Wideband
TDoA	High precision High multi-path resolution	Synchronized anchors Not suitable for mesh network topology	Wideband
RSS	Availability	Poor robustness to fading Model dependent	Wideband Narrowband
AoA	High resolution in LOS	Multiple antennas Sensitivity to NLOS	Narrowband

Table 1. Taxonomy of data acquisition methods

Finally, the third block concerns with the LT algorithm. In table 2, we provide an overview of the approaches typically utilized in the literature.

The first category is based on Bayesian formulations of the LT problem. The Kalman filter (KF) and its variations such as unscendent, cubature and extended are, for instance, the most common Bayesian techniques utilized in the literature. They generally have good performance, although, they are very dependent on system and measurement models assumed in the problem formulation.

The second category of LT methods, instead, is based on non-parametric formulations of the problem. In contrast to Bayesian techniques, non-parametric approaches do not rely on any models and/or assumptions and because of this, the non-parametric methods are widely used in the literature Cox & Cox (2000); Costa et al. (2006); Destino & Abreu (2009); Shang & Ruml (2004); Cheung et al. (2004); Ouyang et al. (2010); Biswas, Liang, Toh, Wang & Ye (2006); Guvenc et al. (2008); Beck et al. (2008). This class of LT techniques can provide very accurate location estimates but they often involve the optimization of a non-convex objective

function. The last category of LT techniques deals with fingerprint methods. The fundamental idea is to search the best pattern match between the stored data and the observations. Fingerprint methods are two-phases approaches. The first one is to construct a database using a priori measurements of the considered parameter (RSS, power profile, ToA, etc.) at different locations, while the second one is to search for the best pattern match between observations and data. This type of LT-technique is very practical in many application scenarios, adapts well with both time- and power-based metrics, but it is sensitive to changes of the environment.

LT Algorithm	Description	Assumption	Weakness
Bayesian	Estimation based on the *a posteriori* pdf	System and measurement models	Dependency on the reliability of the model
Non-parametric	Optimization of on the least-squared error function	None	Typically difficult optimization problem
Fingerprint	Two-phase approach	Map of the physical parameters in the coverage area	Dependency on the accuracy of the map

Table 2. Classification of LT algorithms.

3. Distance-based non-parametric LT system

From now on we focus on distance-based non-parametric LT approaches, where the distance measurements (ranging) are assumed to be the output of either ToA or RSS measurements.

After introducing the basic formulation of the localization problem, we will describe a wavelet based filter to smooth the raw-observations and state-of-the-art optimization methods to minimize the least-squared objective function used in the formulation of the problem.

3.1 LT problem statement

Consider a network deployed in the η-dimensional space and let $\mathbf{X} \in \mathbb{R}^{N\times\eta}$ denote the coordinate matrix, whose i-th row-vector $\mathbf{x}_i \in \mathbb{R}^{\eta}$ indicates the location of the i-th node. The set of indexes $\{i \le N_A\}$ and $\{N_A < i \le N\}$ refer to anchors and targets, respectively. Let $\mathbf{D}$ indicate the Euclidean distance matrix (EDM) of $\mathbf{X}$ obtained from the Euclidean distance function (EDF) $\mathcal{D}(\mathbf{X}) : \mathbb{R}^{N\times\eta} \to \mathbb{R}^{N\times N}$ defined as follows Dattorro (2005)

$$\mathbf{D} = \mathcal{D}(\mathbf{X}) \triangleq \sqrt{\mathbf{1}_N \cdot \mathrm{diag}\left(\mathbf{X}\cdot\mathbf{X}^{\mathrm{T}}\right)^{\mathrm{T}} + \mathrm{diag}\left(\mathbf{X}\cdot\mathbf{X}^{\mathrm{T}}\right)\cdot\mathbf{1}_N^{\mathrm{T}} - 2\cdot\mathbf{X}\cdot\mathbf{X}^{\mathrm{T}}}, \quad (1)$$

where $\mathbf{1}_N$ is a column vector of N elements equal to 1, and diag(·) indicates a column vector containing the diagonal elements of its argument.

The (i, j)-th element of $\mathbf{D}$, denoted by d_{ij}, is the Euclidean distance $\|\mathbf{x}_i - \mathbf{x}_j\|_{\mathrm{F}}$ between the pair of nodes $(\mathbf{x}_i,\mathbf{x}_j)$, where $\|\cdot\|_{\mathrm{F}}$ indicates the Frobenius norm.

The k-th EDM sample $\tilde{\mathbf{D}}_k$ of the set of EDM samples $\{\tilde{\mathbf{D}}_k\}$, are composed of measurements of d_{ij}, denoted by $\tilde{d}_{ij,k}$ described by the ranging model

$$\tilde{d}_{ij,k} = d_{ij} + n_{ij,k}, \tag{2}$$

where $n_{ij,k}$ is assumed to be a Gaussian random variable with zero mean and variance σ_{ij}^2. Let $\mathcal{M}$: $\mathbb{R}^{N\times N\times K} \rightarrow \mathbb{R}^{N\times N}$ denote the functional model of the data-processing block and let $\overline{\mathbf{D}}$ be the smoothed EDM computed as

$$\overline{\mathbf{D}} \triangleq \mathcal{M}(\{\tilde{\mathbf{D}}_k\}), \tag{3}$$

where K is the total number of EDM samples.
Then a non-parametric formulation of the LT problem is given by the following weighted least square (WLS) minimization

$$\begin{aligned} \hat{\mathbf{X}} = \arg\min_{\hat{\mathbf{X}}\in\mathbb{R}^{N\times\eta}} \quad & \frac{1}{2}\cdot\left\|\mathbf{W}\circ\left(\overline{\mathbf{D}}^{\circ q} - \mathcal{D}(\hat{\mathbf{X}})^{\circ q}\right)\right\|_{\mathrm{F}}^2, \\ \text{subject to} \quad & \hat{\mathbf{x}}_i = \mathbf{x}_i \ \forall i = 1,\ldots,N_{\mathrm{A}} \end{aligned} \tag{4}$$

where $\hat{\mathbf{X}} \in \mathbb{R}^{N\times\eta}$ indicates an estimate of $\mathbf{X}$, $\mathbf{W}$ is a weighing matrix that relates to the reliability of $\overline{\mathbf{D}}$, q is an exponent typically chosen amongst the values {1,2} and ○ inidicates the point-wise (Hadamard) power or product, respectively.
The proposed WLS approach is widely used in the literature for several reasons. First, under the assumption of N_{T} = 1 and q = 2, the exact solution of the minimization problem can be computed with a close-form algorithm Beck et al. (2008). Second, under the assumption of low noise, the WLS objective function can be linearized without compromising the accuracy of the location estimates Cheung et al. (2004); Guvenc et al. (2008). Third, under the assumption of a zero-mean Gaussian noise and q = 1, the WLS approach is equivalent to the maximum-likelihood (ML) formulation of the LT problem Patwari et al. (2003); Biswas, Liang, Toh & Wang (2006). Indeed, by computing the likelihood function of $\hat{\mathbf{X}}$

$$p(\hat{\mathbf{X}} \mid \overline{\mathbf{D}}) = \frac{1}{(2\pi)^{\eta}} \prod_{e_{ij}\in E} \exp\left(-\frac{(\overline{d}_{ij} - \| \hat{\mathbf{x}}_i - \hat{\mathbf{x}})_j \|_2)^2}{\sigma_{ij}^2 / K_{ij}}\right), \tag{5}$$

where K_{ij} is the number of measurements of d_{ij}, e_{ij} indicates an connected link between the ith and the j-th nodes, and E as the set of all connected link, and taking the logarithm it follows that

$$\ln\left(p(\hat{\mathbf{X}} \mid \overline{\mathbf{D}})\right) = \frac{1}{(2\pi)^{\eta}} \sum_{e_{ij}\in E} \frac{K_{ij}}{\sigma_{ij}^2}\cdot\left(\overline{d}_{ij} - \| \hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j) \|_2\right)^2, \tag{6}$$

which is equivalent to equation 4 rewritten as

$$\sum_{e_{ij}\in E} w_{ij}^2 \cdot \left(\overline{d}_{ij}^{q} - \| \hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j) \|_2^q\right)^2, \tag{7}$$

with $w_{ij}^2 = \frac{K_{ij}}{\sigma_{ij}^2}$ and q = 1.

4. Ranging post-processing

As shown in figure 2, while block-1 deals with the detection, acquisition and association problem, block-2 pre-filters the data to improve the signal to noise ratio for the measurements. Although several algorithms can be used to smooth the observations, *e.g.* moving average, exponential, autoregressive moving average, Kalman filters and so forth, here after we will focus on a low-complex wavelet-based pre-filtering that has been proved to suit the localization and target tracking scenarios.

4.1 Wavelet-based smoothing

The Wavelet Transform (WT) makes use of a unique dilated window (the wavelet function) to analyze signals. This allows good time resolution (for short windows) at high frequency and good frequency resolution (corresponding to long-window) at low frequency S.Mallat (1998). The decomposition is based on a family of functions $\sqrt{s}\psi(s(x-u))_{(s,u)\in\mathbb{R}^2}$ corresponding to the translated and dilated version of the *wavelet* function $\psi(x)$, also called *mother* wavelet, and with s and u corresponding to the *scaling* and *translation* factors.

Given a continuous function $f(x)$, its continuous wavelet transform, here denoted as $Wf(s,u)$, corresponds to the inner product $\langle f(x),\psi_s(x-u)\rangle$, meaning the cross-correlation between the original function and the scaled wavelet shifted at u. In S.Mallat & S.Zhong (1992) it is shown that choosing $\psi(x) = d\phi(x)/dx$, with $\phi(x)$ as smoothing function, then it is possible to characterize the shape of irregular functions $f(x)$ through $Wf(s,u)$. In addition, using the properties of the convolution operator it follows that

$$Wf_s(x) = f * \left(s\frac{d\phi_s}{dx} \right)(x) = s\frac{d}{dx}\left(f * \phi_s \right)(x), \tag{8}$$

which allows to interpret $Wf(s,u)$ as the derivative of a local average of $f(x)$, with smoothing degree depending on the scale factor s.

Amongst the several algorithm to compute the wavelet transform of discrete signals, because of its low complexity and its redundant representation of the signal $f(x)$ across the scales, which has been proved to be particularly suitable in filtering applications, we use the *á trous* algorithm briefly summarized in the following.

Let $a_0[n]$ be the discrete signal to be analyzed, with n as the discrete time index and assume the value for $a_0[n]$ in n equivalent to the local average between the original continuos function $f(x)$ and a kernel function $\phi(x-n)$ (namely $\langle f(x),\phi(x-n)\rangle$), then at any scale $j > 0$, a smoothed version of $a_0[n]$ is computed as $\langle f(x),\phi_{2^j}(x-n)\rangle$, with

$$\phi_{2^j}(t) = \frac{1}{\sqrt{2^j}}\phi\left(\frac{x}{2^j} \right). \tag{9}$$

The function $\phi(x)$ is called *scaling function* and it corresponds to a low-pass filter, while the coefficient for the dyadic WT are obtains by $z_{2^j}[n] = Wf(s,n) = \langle f(x),\psi_{2^j}(x-n)\rangle$ with $\psi_{2^j}(x-n)$ defined similarly to equation 9. Once the low-pass filter $h[n]$ and high-pass filter $g[n]$ are designed then $a_0[n]$ is decomposed by repetitively computing

$$a_{j+1}[n] = a_j * \tilde{h}_j[n], \\ d_{j+1}[n] = a_j * \tilde{g}_j[n], \tag{10}$$

with h_j $[n]$ obtained from h $[n]$ inserting 2^j –1 zeros between each sample of the filter (similarly for g_j $[n]$) and $0 < j < J$.

We use the WT to study each time series corresponding to subsequent ranging measured at the devices. We restrict ourselves to LoS target tracking scenarios, and we suggest a scheme to adaptively pre-filtering the observations f $[n]$ in a completely non-parametric fashion. To do so we use the output of the DWT to estimate σ_d and δ, namely the noise level affecting the observations and the target dynamic perceived at each anchor via ranging.

As mentioned above, the wavelet coefficients d_j $[n]$ computed at the different scales j include the high frequency components for the original signal and it is therefore used to characterize σ_d. Similarly, the output of the scaling function is used to infer δ. Clearly the approach works at best if the signal can be decomposed in high frequency components of short duration and a low frequency part of relatively long duration. From equations 8 and 10 it is clear that the wavelet coefficients $d_j[n]$ at the scale j = 1 represents a simple differential operator, therefore under static ($\hat{v}$ = 0) or anyway scenario characterized by a small dynamic $\hat{v}$, an estimate of $\hat{\sigma}_d$ can be inferred from $d_1[n]$.

However, when the target dynamic ($\hat{v}$) increases, $d_1[n]$ starts including part of the energy associated to f $[n]$ and eventual estimates of $\hat{\sigma}_d$ would be affected by error. To overcome this problem $\hat{\sigma}_d$ is estimated as the standard deviation of $d_1[n]$ computed from subsets of subsequent observations characterized by the same polarity value in the support function

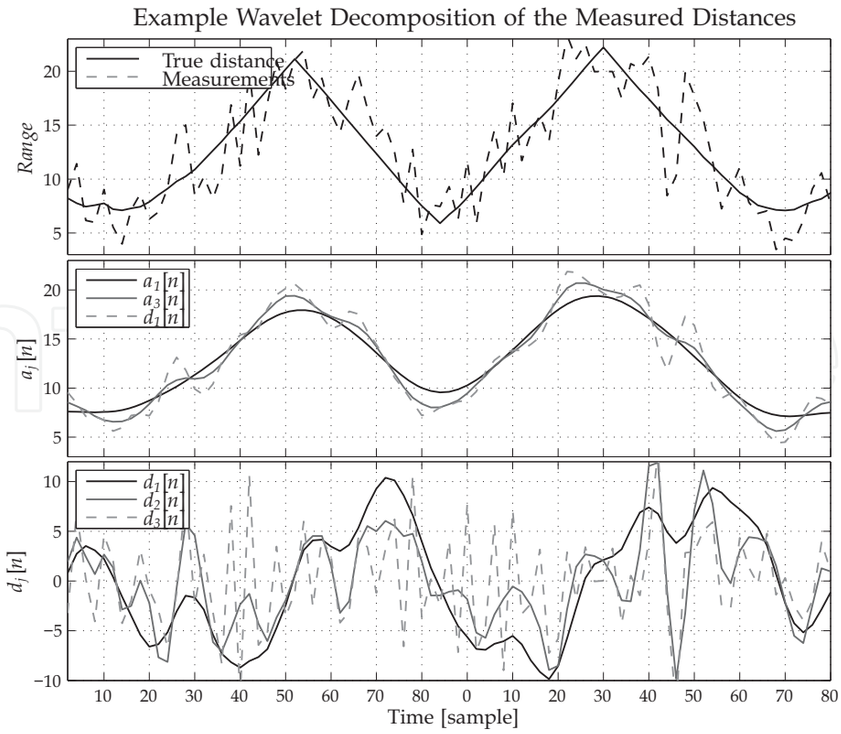


Fig. 3. Example wavelet decomposition of TOA ranging.

described in section 5.1.2. To distinguish the long term time process associated to the real TOA measurements (low-pass filtered version of $f[n]$), we use an averaged version of $a_1[n]$ in the same way proposed in Macagnano & de Abreu (2008), meaning that at each sampling time we compute the DWT on a window of size 2^J and centered at n.
From this averaged $a_1[n]$ we compute the parameter δ approximating the perceived dynamic at the specific anchor with respect to the considered target. The decomposition of $f[n]$ in its high/low frequency components is performed subject to the boolean operator Θ defined in section 5.1.2.
Using Θ, computed at each time n and for each anchor-to-target link, we decide whether the real ToA observation is better approximated by the measured ranging ($f[n]$) or it low-pass filtered version ($a_1[n]$). The only price paid using this wavelet smoothing based on Θ, is the introduction of a lag of $2^J - 1$ samples in the computation.

5. LT algorithm

The LT problem formulation considered in this chapter is the WLS-ML approach

$$\begin{aligned} \hat{\mathbf{X}} = \arg\min_{\hat{\mathbf{X}} \in \mathbb{R}^{N \times \eta}} & \sum_{e_{ij} \in E} w_{ij}^2 \cdot \left(\bar{d}_{ij} - \| \hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j \|_{\mathrm{F}} \right)^2 \\ \text{subject to} \quad & \hat{\mathbf{x}}_i = \mathbf{x}_i \ \forall i = 1, \ldots, N_{\mathrm{A}}. \end{aligned} \tag{11}$$

The challenges faced in this optimization problem are: the computation of the weights and the minimization of the objective function. In the sequel, we will tackle both issues and we will describe in details very effective solutions.

5.1 Weighing strategies

In the optimization problem posed in equation 11, the purpose of the weights is "*to reflect differing levels of concern about the sizes of the terms*" in the objective function. In other words, higher the weight tighter is the concern Boyd & Vandenberghe (2004).
The first proposed strategy, nevertheless optimal in the ML sense, can be derived directly from equation 6,

$$w_{ij}^* = \frac{K_{ij}}{\sigma_{ij}^2}, \ \forall \sigma_{ij}^2 \neq 0. \tag{12}$$

For the special case of $\sigma_{ij}^2 = 0$, *i.e.* no error, $w_{ij}^* = 0$ but in equation 11, we add the equality constraint

$$\hat{d}_{ij} = \tilde{d}_{ij}. \tag{13}$$

In most cases, however, $\sigma_{ij}^{2\prime}$s are not known *a priori*, therefore, alternative weighing strategies will be considered. The first alternative referred to as binary weight or unweighted is

$$w_{ij}^u = \begin{cases} K_{ij}, & \forall e_{ij} \in E, \\ 0, & otherwise \end{cases}. \tag{14}$$

This method is also optimal if $\sigma_{ij}^2 = \sigma^2 \ \forall ij$ (*ij* because $\frac{1}{\sigma}$ becomes a constant and therefore, it is a common factor to the WLS-ML objective function.
The second alternative is to replace in equation 12 σ_{ij}^2 by the sample variance $\hat{\sigma}_{ij}^2$ estimated either in the wavelet filter as described in section 5.1.2 or computed as

$$\hat{\sigma}_{ij}^2 = \frac{1}{K_{ij}-1}\sum_{k=1}^{K_{ij}}\left(\tilde{d}_{ij,k} - \check{d}_{ij}\right)^2, \tag{15}$$

where $\check{d}_{ij}$ is the sample mean computed as

$$\check{d}_{ij} = \frac{1}{K_{ij}}\sum_{k=1}^{K_{ij}}\tilde{d}_{ij,k}. \tag{16}$$

A third method is based on the regression model proposed in Costa et al. (2006). This technique is very effective only if a sufficient number of measurements are collected and if short distances are more reliable than long ones. In this case, weights are given by

$$w_{ij}^e = \sum_{k=1}^{K_{ij}} \exp\left(\frac{-\tilde{d}_{ij,k}^2}{\left(\max_j \tilde{d}_{ij,k}\right)^2}\right), \forall e_{ij} \in E. \tag{17}$$

Yet another alternative weighing strategy, which will be described with more details in the following subsection, is based on the relationship between the concept of concern (seen as constraint in the optimization) and the notion of statistical confidence Destino & De Abreu (2009). In essence, the weight, hereafter referred to as dispersion weight, will be computed as "*a measure of the confidence about the estimate* $\bar{d}_{ij}$ *associated with the penalty on the assumption of LOS conditions*".

5.1.1 Dispersion weights

The dispersion weight is mathematically formulated as

$$w_{ij}^D \triangleq Pr\left\{-\gamma \le \varepsilon_{ij} \le \gamma\right\} \cdot \mathcal{P}_{ij}, \tag{18}$$

where γ is the confidence bound of the observation $\bar{d}_{ij}$ around the true distance d_{ij}, $\mathcal{P}_{ij}$ is a penalty imposed over the LOS assumption and $\varepsilon_{ij} \triangleq d_{ij} - \bar{d}_{ij}$.
For convenience, we shall hereafter use w_{ij}^{L} in reference to the probability in equation 18, such that $w_{ij} = w_{ij}^L \cdot \mathcal{P}$. The weights w_{ij}^L and $\mathcal{P}_{ij}$ will also be dubbed the confidence weights and penalty weights, respectively.
Under the assumptions that $\rho_{ij} = 0$, *i.e.* LOS channel conditions, and $\tilde{d}_{ij,k}$ are independent, the dispersion weight can be rewritten in the form

$$w_{ij}^{\mathrm{L}} = 2 \cdot \Pr\left\{\varepsilon_{ij} \le \gamma\right\} - 1, \tag{19}$$

where we use the LOS assumption to set $\mathcal{P}_{ij} = 1$.

Considering $\tilde{d}_{ij,k}$ as Gaussian random variable, by means of small-scale statistics w_{ij}^{L} can be computed as Gibbons (1992),

$$w_{ij}^{\mathrm{L}}(\hat{\sigma}_{ij}, K_{ij}; \gamma) = -1+2 \cdot \int_{-\infty}^{T_{ij}} f_{\mathrm{T}}(t; K_{ij}-1)\, dt, \tag{20}$$

$$T_{ij} = \gamma \cdot \sqrt{K_{ij} / \hat{\sigma}_{ij}^2}, \tag{21}$$

where $f_{\mathrm{T}}(t;n)$ is the T-distribution of n degree of freedom and T_{ij} is the *t-score*.
As emphasized by the notation, w_{ij}^{L} is a function of the sample variance $\hat{\sigma}_{ij}^2$ and the number of samples K_{ij}, as well as the confidence bound γ, to be specified below. Since $\hat{\sigma}_{ij}^2$ and K_{ij} carry different information about the true value of d_{ij}, it is not surprising that both these parameters impact on the weight w_{ij}^{L}. In fact, the plots of $w_{ij}^{\mathrm{L}}(\hat{\sigma}_{ij}, K_{ij}; \gamma)$ illustrated in figures 4 and 5, show that w_{ij}^{L} grows with the inverse of $\hat{\sigma}_{ij}^2$ (for fixed K_{ij}), and with K_{ij} (for fixed $\hat{\sigma}_{ij}^2$). This is in accordance with the argument outlined in the heading of this section and widely invoked by other authors Biswas, Liang, Toh, Wang & Ye (2006); Costa et al. (2006); Shang & Ruml (2004); Patwari et al. (2003); Boyd & Vandenberghe (2004); Alfakih et al. (1999), since $\hat{\sigma}_{ij}^2$ is proportional to the uncertainty of $\bar{d}_{ij}$, as a measure of d_{ij}, while K_{ij} relates to the quality of $\bar{d}_{ij}$ and $\hat{\sigma}_{ij}^2$ as measures of d_{ij} and its dispersion, respectively.
Unlike $\hat{\sigma}_{ij}^2$ and K_{ij}, which are obtained in the process of measuring inter-node distances, the confidence bound γ is a free choice parameter that allows for fine-tuning the relative values of w_{ij}^{L}.

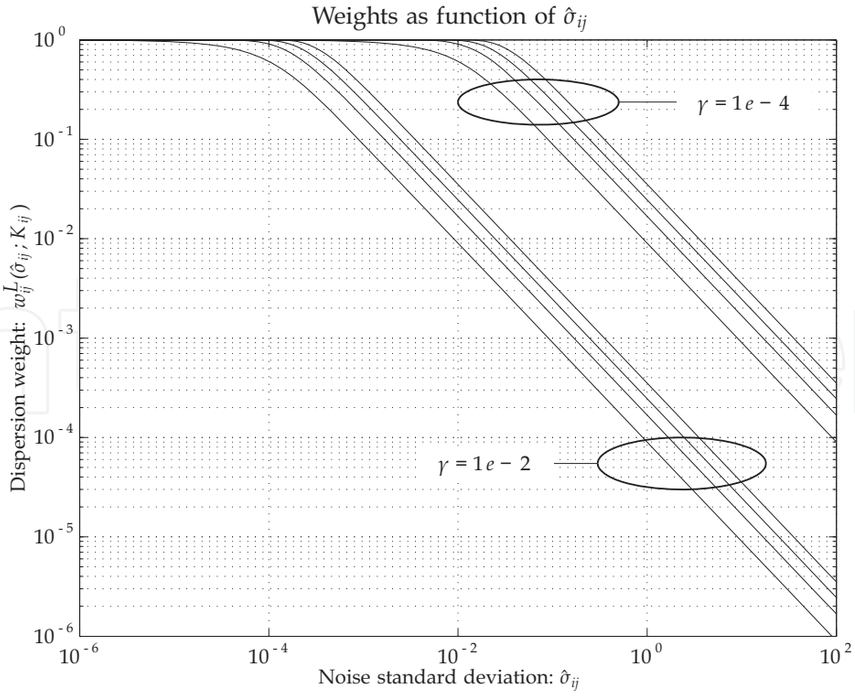


Fig. 4. Dispersion weight as functions of $\hat{\sigma}_{ij}$

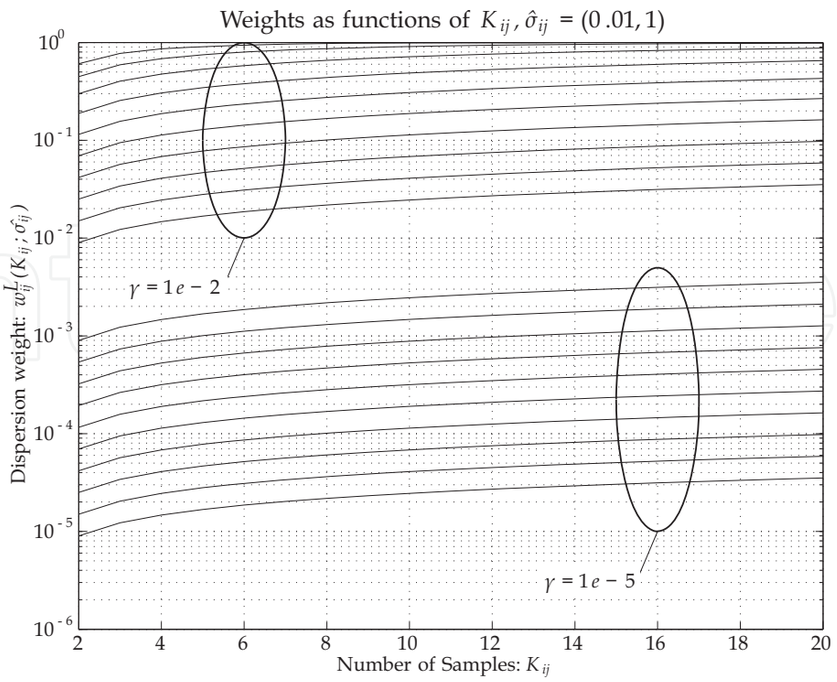


Fig. 5. Dispersion weight as functions of K_{ij}

The mechanism to find the optimum γ is given by the following optimization problem based on the *diversity index* or *entropy* metric

$$\gamma_{\text{opt}} = \arg\max_{\gamma \in \mathbb{R}^{+}} \mathcal{H}(\gamma), \tag{22}$$

where

$$\mathcal{H}(\gamma) = -\sum_{k=K_{\min}}^{K_{\max}} \int_{\sigma_{\min}}^{\sigma_{\max}} w^{L}(s,k;\gamma) \cdot \ln\left(w^{L}(s,k;\gamma)\right) ds, \tag{23}$$

where $K_{\min}$, $K_{\max}$, $\sigma_{\min}$ and $\sigma_{\max}$ are the minimum and the maximum number of observable samples and the minimum and the maximum typical ranging error, respectively.
The derivation of the method is omitted in this book, however, an interested reader can refer to Destino & De Abreu (2009). To validate the aforementioned optimization criterion, in figure 6 we show that varying γ, the minimum root-mean-square-error obtained via solving the optimization in equation 11 is close to that one achieved with γ_{opt}.
For the sake of completeness, in figure 7 we illustrates γ_{opt} as function of $K_{\max}$ and $\sigma_{\max}$, considering $K_{\min} = 2$, $\sigma_{\min} = 1e\text{-}4 \approx 0$. The same results are also shown in table 3.

5.1.2 Dynamic weighing strategy

In this section it is shown how to use the output of the wavelet transform of the time series $f[n]$ corresponding to the ToA measurements at each anchor node to extract a confidence on the observations in the form of

$$w_i[n] = \frac{1}{\hat{\sigma}_d}. \tag{24}$$

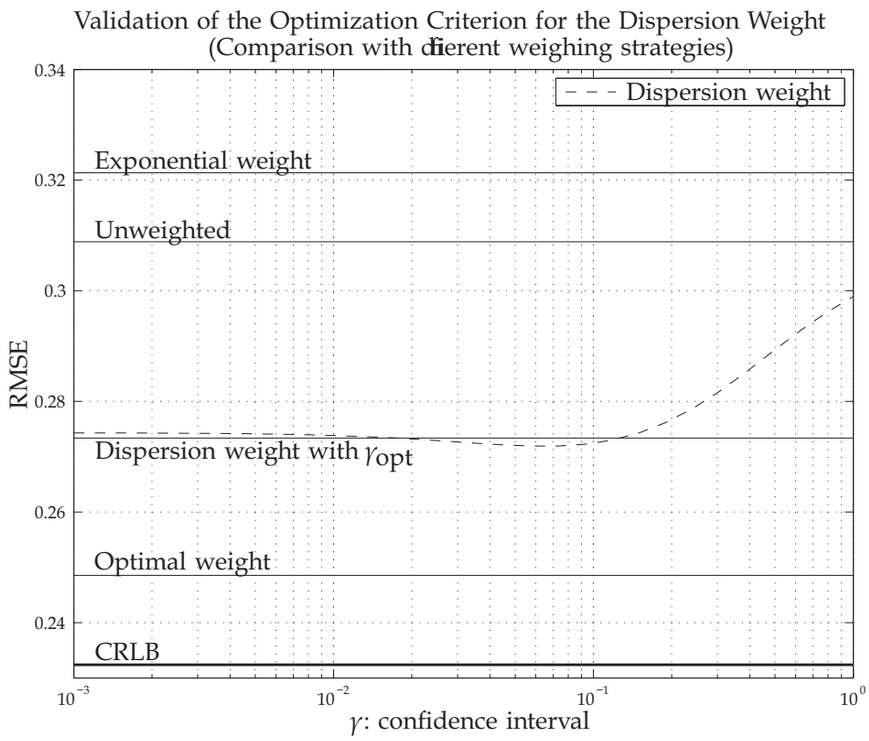


Fig. 6. Validation of the optimization criterion of γ. Simulation results are obtained for a network with $N_A = 4$, $N_T = 1$, $\sigma_{min} =$ 1e-4, $\sigma_{max} = 2$, $K_{min} = 2$ and $K_{max} = 7$.

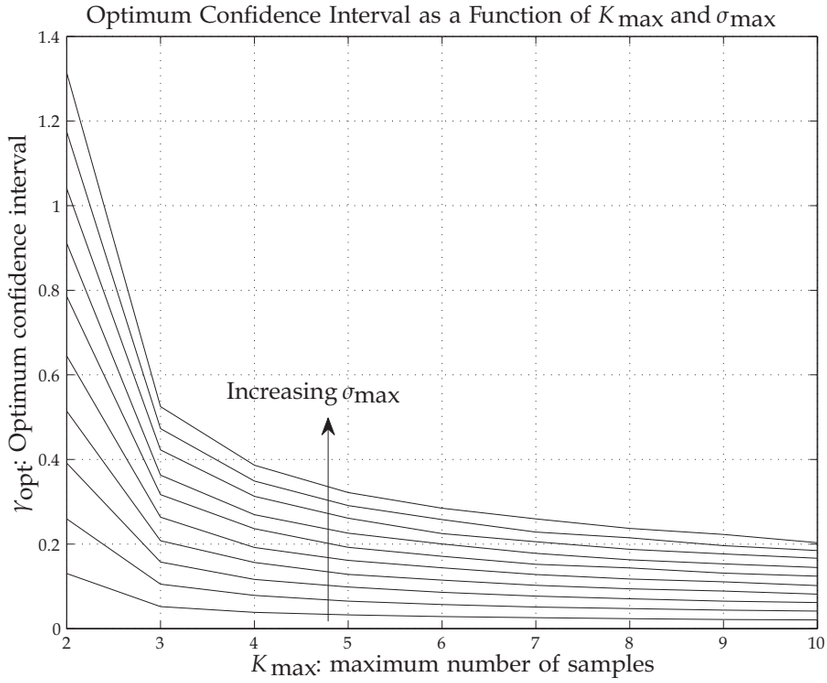


Fig. 7. Plot of the optimal confidence interval as a function of σ_{max} and K_{max}, with $\sigma_{min} =$ 1e-4 and $K_{min} = 2$.

K_{max}	σ_{max}									
	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80	2.00
2	13.04	25.94	39.13	51.41	64.49	78.58	91.10	104.04	117.45	131.39
3	5.20	10.51	15.75	20.79	26.36	31.67	36.28	42.25	47.28	52.48
4	3.86	7.85	11.63	15.66	19.19	23.62	26.96	31.29	34.91	38.65
5	3.24	6.52	9.85	12.85	16.14	19.24	22.56	26.13	29.10	32.16
6	2.86	5.70	8.58	11.47	14.38	17.11	20.03	22.49	25.78	28.46
7	2.59	5.13	7.69	10.24	12.81	15.22	17.79	20.55	22.83	25.96
8	2.35	4.74	7.08	9.41	11.75	14.35	16.27	18.78	21.49	23.69
9	2.18	4.37	6.51	8.89	11.09	13.14	15.33	17.68	19.62	22.28
10	2.07	4.15	6.16	8.17	10.17	12.39	14.45	16.65	18.46	20.33

Table 3. Tabulation of the optimal confidence interval γ_{opt}. Values are indicated with the multiplicative factor 1e-2.

Obviously the problem is to recover a sufficiently good estimate of $\hat{\sigma}_d$ from the time varying process $f[n]$. To do so we use a boolean support to distinguish, amongst subsequent observations of $f[n]$, the one belonging to the same trend. To do it, we check whether the coefficient $d_j[n]$ at n and computed as accordingly to subsection 4.1, retain the same polarity across the scales $2 < j < J$. Whether the condition above is satisfied, then we assign to the nth value of our support ±1, accordingly to the polarity of $d_1[n]$, differently we assign 0. The reason to distinguish the oscillatory behavior of $f[n]$ using the multiresolution representation provided by $d_j[n]$ $(2 < j < J)$ is that it better characterizes the trend of $f[n]$, since it allows to assign 0 to the points corresponding to significantly high discontinuity on $f[n]$, such as direction changes in the targets trajectories. Thus, computing the support above at each time n and of each anchor-to-target link, it is possible to distinguish amongst subsequent measurements, the one belonging to the same set of subsequent points in the support, and therefore mainly affected by the energy of the noise process.
Consequently $\hat{\sigma}_d$ is estimated as the standard deviation for $d_1[n]$, from those subsets of subsequent observations characterized by the same polarity value in the support. In case of the support equal to 0 at n, then we simply keep the previous estimate. However, since $d_1[t]$ is computed convolving $f[t]$ with $g_1[n]$, then the estimated $\hat{\sigma}_d$ described in section 4.1needs to be compensated accordingly to the values of $g[n]$, which by $\hat{\sigma}_j^e = \hat{\sigma}_d \sigma_j$, with $\hat{\sigma}_j^e$ as the value estimated directly from $d_j[t]$ and σ_j a reference value pre-computed and dependent on the wavelet used, namely the filter $g[n]$. Therefore $\hat{\sigma}_d$ is computed from $d_1[n]$ from the points belonging to the same subsets on the support as

$$\hat{\sigma}_d = \hat{\sigma}_j^e / \sigma_j. \tag{25}$$

A limit on the minimum number points representing the subsets has to be fixed. In our simulation this value was fixed to 3. To control whether $f[n]$ can be decomposed in high frequency components of short duration and a low frequency part of relatively long duration, meaning that we test whether if is feasible to approximate $f[n]$ using its low-pass filtered version ($a_1[n]$), we use the boolean operator Θ defined by

$$\begin{cases} \delta > \sigma_d \rightarrow \Theta = 1, \\ \delta < \sigma_d \rightarrow \Theta = 0. \end{cases} \tag{26}$$

5.2 Optimization methods

After the calculation of the weights, the next challenge is to minimize the objective function in equation 11. In the sequel, we will describe state-of-the-art as well as novel optimization methods to compute reliable solutions.

5.2.1 Classical-multidimensional scaling

The classical-multidimensional scaling (CMDS) approach Cox & Cox (2000) can be thought of as an algebraic solution of the ML-WLS localization problem with $\mathbf{W}=\mathbf{1}_N \cdot \mathbf{1}_N^{\mathrm{T}} - \mathbf{I}_N$, where $\mathbf{I}_N \in \mathbb{R}^{N\times N}$ indicates the identity matrix. Under the assumption that all d_{ij} are measured at least once ($K_{ij} \neq 0 \forall (i, j)$), this solution is the least squares solution of equation 4 Cox & Cox (2000). The fact that no iterations are required makes the CMDS a very low-complexity and fast solution of the localization problem. It can be shown, however, that in the presence of incomplete EDM samples, the performance of the CMDS algorithm degrades drastically. Therefore, in the context of incomplete mesh network topology, the CMDS or some its variations, will be used to initialize iterative minimization techniques.

Briefly, the CMDS algorithm can be summarized as follows. First, compute the kernel $\overline{\mathbf{K}}$

$$\overline{\mathbf{K}} \triangleq \mathcal{K}(\overline{\mathbf{D}}) = -(\mathbf{T}\cdot(\overline{\mathbf{D}})^{\circ 2}\cdot\mathbf{T})/2, \tag{27a}$$

$$\mathbf{T} \triangleq \mathbf{I}_N - (\mathbf{1}_N \cdot \mathbf{1}_N^{\mathrm{T}})/N. \tag{27b}$$

Then, an estimate of the node coordinates $\hat{\mathbf{X}}$ is obtained as

$$\hat{\mathbf{X}} = \left([\mathbf{U}]_{\mathrm{UL}:N\times\eta} \cdot [(\Lambda)^{\frac{1}{2}}]_{\mathrm{UL}:\eta\times\eta} \right)^{T}, \tag{28}$$

where $[\cdot]_{\mathrm{UL}:n\times q}$ denotes the n-by-q upper-left partition and the matrices $\mathbf{U}$ and Λ are the eigenvector and eigenvalue matrices of $\overline{\mathbf{K}}$ Cox & Cox (2000), respectively, both in decreasing order.

5.2.2 Nyström algortihm

An alternative to the CMDS is the *Nyström approximation*[1] technique Williams & M.Seeger (2000); C. Fowlkes & Malik (2004). This method performs the same eigen-decomposition of CMDS, but in a more efficient manner.

Consider the Nyström kernel given by C. Fowlkes & Malik (2004)

$$\tilde{\mathbf{K}} \approx \left[\begin{array}{c|c} [\mathbf{K}]_{1:\eta,1:\eta} & [\mathbf{K}]_{1:\eta,\eta+1:N} \\ \hline [\mathbf{K}]_{1:\eta,\eta+1:N}^{\mathrm{T}} & [\mathbf{K}]_{1:\eta,\eta+1:N}^{\mathrm{T}} \cdot [\mathbf{K}]_{1:\eta,1:\eta}^{-1} \cdot [\mathbf{K}]_{1:\eta,\eta+1:N} \end{array}\right], \tag{29}$$

in which $[\mathbf{K}]_{1:\eta,1:\eta}$ and $[\mathbf{K}]_{1:\eta,\eta+1:N}$ denote the upper-left η-by-η, and the upper-right η-by-$(N-\eta)$ minors of $\mathbf{K}$.

Recall that (Dattorro, 2005, pp. 195)

[1] In the case of Euclidean kernels, the Nyström "approximation" is actually an exact completion if the entries of the required minors are error-free.

$$[\mathbf{K}]_{1:\eta,1:\eta} = -\frac{1}{2}\left([\mathbf{D}]_{1:\eta,1:\eta} + \mathbf{C}_1 \otimes \mathbf{1}_\eta \mathbf{1}_\eta^{\mathrm{T}} - \mathbf{C}_2 \otimes \mathbf{1}_\eta^{\mathrm{T}} - \mathbf{C}_3 \otimes \mathbf{1}_\eta\right), \tag{30}$$

$$[\mathbf{K}]_{1:\eta,\eta+1:N} = -\frac{1}{2}\left([\mathbf{D}]_{1:\eta,\eta+1:N} + \mathbf{C}_1 \otimes \mathbf{1}_\eta \mathbf{1}_{N-\eta}^{\mathrm{T}} - \mathbf{C}_2 \otimes \mathbf{1}_{N-\eta}^{\mathrm{T}} - \mathbf{C}_4 \otimes \mathbf{1}_\eta\right), \tag{31}$$

where $\otimes$ denotes the Kronecker product and

$$\mathbf{C}_1 = \frac{1}{\eta^2} \cdot \left[\mathbf{1}_\eta^{\mathrm{T}} \cdot [\mathbf{D}]_{1:\eta,1:\eta} \cdot \mathbf{1}_\eta\right], \tag{32}$$

$$\mathbf{C}_2 = \frac{1}{\eta} \cdot \left[[\mathbf{D}]_{1:\eta,1:\eta} \cdot \mathbf{1}_\eta\right], \tag{33}$$

$$\mathbf{C}_3 = \frac{1}{\eta} \cdot \left[\mathbf{1}_\eta^{\mathrm{T}} \cdot [\mathbf{D}]_{1:\eta,1:\eta}\right], \tag{34}$$

$$\mathbf{C}_4 = \frac{1}{\eta} \cdot \left[\mathbf{1}_\eta^{\mathrm{T}} \cdot [\mathbf{D}]_{1:\eta,\eta+1:N}\right]. \tag{35}$$

Finally, invoke the relation (Dattorro, 2005, pp. 196)

$$\tilde{\mathbf{D}} = \left(\mathbf{1}_N \cdot \mathrm{diag}\left(\tilde{\mathbf{K}}\right)^{\mathrm{T}} + \mathrm{diag}\left(\tilde{\mathbf{K}}\right) \cdot \mathbf{1}_N^{\mathrm{T}} - 2 \cdot \tilde{\mathbf{K}}\right). \tag{36}$$

Equation 36 yields a complete set of distances associated with $\tilde{\mathbf{K}}$, such that any missing entries of $[\mathbf{D}]_{\eta+1:N,\eta+1:N}$ can be replaced by corresponding entries from $\tilde{\mathbf{D}}$.

At this point, let us emphasize that $[\mathbf{D}]_{1:\eta,1:\eta}$ contains the distances amongst anchors and consequently $[\mathbf{K}]_{1:\eta,1:\eta}$, $\mathbf{C}_1$, $\mathbf{C}_2$ and $\mathbf{C}_3$ are all *constant*, such that $\tilde{\mathbf{K}}$ can be updated very efficiently.

Furthermore, the elements of $[\mathbf{D}]_{1:\eta,\eta+1:N}$ are the distances from anchors to targets, and therefore constitute the least (reasonable) amount of information required by tracking applications, such that this "completion" procedure can always[2] be applied.

In the extreme case of $[\mathbf{D}]_{\eta+1:N,\eta+1:N} = \mathbf{0}_{N-\eta}$, then to recover $[\mathbf{X}]_{\eta+1:N,1:\eta}$ only the eigendecomposition of $[\tilde{\mathbf{K}}]_{1:\eta,1:\eta}$ is required C. Fowlkes & Malik (2004).

Indeed, let $[\tilde{\mathbf{K}}]_{1:\eta,1:\eta} = \mathbf{Q} \cdot \mathbf{\Lambda} \cdot \mathbf{Q}_{\mathrm{T}}$ as the eigendecomposition of $[\tilde{\mathbf{K}}]_{1:\eta,1:\eta}$. From equation 28 it follows that $[\mathbf{X}]_{1:\eta,1:\eta} = \mathbf{Q} \cdot \mathbf{\Lambda}^{\frac{1}{2}}$, and because $[\tilde{\mathbf{K}}]_{1:\eta,\eta+1:N} = [\mathbf{X}]_{1:\eta,1:\eta} \cdot [\mathbf{X}]_{\eta+1:N,1:\eta}^{\mathrm{T}}$ then

$$[\mathbf{X}]_{\eta+1:N,:} = [\mathbf{X}]_{1:\eta,1:\eta}^{-\mathrm{T}} \cdot [\tilde{\mathbf{K}}]_{1:\eta,\eta+1:N} = \mathbf{Q} \cdot \Lambda^{\frac{1}{2}} \cdot [\tilde{\mathbf{K}}]_{1:\eta,\eta+1:N}, \tag{37}$$

In conclusion, if an incomplete EDM is observed the aforementioned steps can be followed to complete **D** before constructing the MDS kernel **K*** described in equations 27a.

---

[2] Even the case of sparse incomplete EDMs in which *none* of the rows of **D** is complete could, in principle, also be dealt with by combining the Nyström solution with standard completion algorithms applied to a restricted subset of η rows of **D** Shang & Ruml (2004). This case, however, is of relatively little interest to tracking applications and outside the scope of the article.

5.2.3 SMACOF

The SMACOF technique is a well-known iterative algorithm that attempts to find the minimum of a non-convex function by tracking the global minima of the so-called majored convex functions $\mathcal{T}(\hat{\mathbf{X}},\mathbf{Y})$ successively constructed from the original objective and basis on the previous solutions. In our context, the objective function to majorize is that one given in equation 7. Thus, we have

$$\mathcal{T}(\hat{\mathbf{X}},\mathbf{Y}) = \sum w_{ij}^2 \cdot \overline{d}_{ij}^{\,2} + \mathrm{tr}\left(\hat{\mathbf{X}}^{\mathrm{T}} \cdot \mathbf{H} \cdot \hat{\mathbf{X}}\right) - 2 \cdot \mathrm{tr}\left(\hat{\mathbf{X}}^{\mathrm{T}} \cdot \mathbf{A}(\mathbf{Y}) \cdot \mathbf{Y}\right), \tag{38}$$

where tr(·) denotes the trace, $\mathbf{Y} \in \mathbb{R}^{N\times\eta}$ is an auxiliary variable and the entries of $\mathbf{H}$ and $\mathbf{A}(\mathbf{Y})$ are given by

$$h_{ij} = \begin{cases} \sum_{\substack{i=1 \\ i\neq j}}^{N} h_{ij}, i = j, \\ -w_{ij}^2, i \neq j, \end{cases} \tag{39a}$$

$$a_{ij} = \begin{cases} \sum_{\substack{i=1 \\ i\neq j}}^{N} a_{ij}, i = j, \\ w_{ij}^2 \cdot \dfrac{\overline{d}_{ij}}{\| \mathbf{y}_i - \mathbf{y}_j \|_2}, i \neq j, \end{cases} \tag{39b}$$

where $w_{ij} > 0$ if $e_{ij} \in E$ and $w_{ij} = 0$ otherwise.

At the ℓ-th iteration the global minimum $\hat{\mathbf{X}}^{(\ell)}$ of the majored function $\mathcal{T}(\hat{\mathbf{X}},\mathbf{Y})$ with $\mathbf{Y} = \hat{\mathbf{X}}^{(\ell-1)}$, is computed via the Guttman transform,

$$\hat{\mathbf{X}}^{(\ell)} = \mathbf{H}^{\dagger} \cdot \mathbf{A}\left(\hat{\mathbf{X}}^{(\ell-1)}\right) \cdot \hat{\mathbf{X}}^{(n-1)}, \tag{40}$$

where † denotes the pseudoinverse.

5.2.4 Linear global distance continuation

While the C-MDS and the Nyström approximation are algebraic approaches and SMACOF relies on the initialization point $\hat{\mathbf{X}}^{(0)}$, the algorithm proposed below, performs a low-complexity unconstrained global optimization. The approach is based on an iterative global smoothing technique, in which the global minimum is sought (with probability close to 1) after L number of iterations. The overall worse-case complexity of the method is equal to $L \times O$, where O is the worse-case complexity of the optimization technique used at the ℓ-th iteration. For convenience, we use a Quasi-Newton line search method whose complexity is $\mathcal{O}(N^2)$ Nocedal & Wright (2006). The proposed technique will be hereafter referred to as linear-global distance continuation (L-GDC) method, inspired by More & Wu (1997).

5.2.4.1 Fundamentals of the L-GDC

The objective of this subsection is to provide the fundamental Definitions, Theorems and Lemmas used in the L-GDC algorithm. Given the limited number of pages, we omit all proofs which can be found in Destino & Abreu (2010); More & Wu (1997).

Definition 1 (Gaussian kernel) *Let $g(u,\lambda)$ be the Gaussian kernel defined as*

$$g(u,\lambda) \triangleq e^{-u^2/\lambda^2}. \tag{41}$$

Definition 2 (Gaussian transform) *Let $\langle s\rangle_\lambda(\mathbf{x})$ denote the Gaussian transform (smoothed function) of a function $s(\mathbf{x})$, and given by*

$$\langle s\rangle_\lambda(\mathbf{x}) \triangleq \frac{1}{\pi^{n/2}\lambda^n}\int_{\mathbb{R}^n} s(\mathbf{u}) e^{-\frac{\|\mathbf{x}-\mathbf{u}\|_{\mathrm{F}}^2}{\lambda^2}} d\mathbf{u}, \tag{42}$$

where $\mathbf{u},\mathbf{x} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}^+$ is a parameter that controls the degree of smoothing ($\lambda \gg 0$ strong smoothing).

Theorem 1 (Continuation method) *Let $\{\lambda^{(\ell)}\}$ with $\{1 \le \ell \le L\}$ be any sequence of λ's converging to zero, i.e. $\lambda^{(L)} = 0$. If $\mathbf{x}^{(\ell)}$ is a global minimizer of $\langle s\rangle_{\lambda^{(\ell)}}(\mathbf{x})$ and $\{\mathbf{x}^{(\ell)}\}$ converges to $\mathbf{x}^*$, then $\mathbf{x}^*$ is a global minimizer of $\langle s\rangle_{\lambda^{(\ell)}}(\mathbf{x})$.*

Theorem 2 (WLS-ML Smoothed Function) *Let $s(\hat{\mathbf{X}})$ equal to the objective function in equation 7, and for simplicity consider $\eta = 2$. Then, the smoothed function $\langle s\rangle_\lambda(\hat{\mathbf{X}})$ is given by*

$$\begin{aligned}\langle s\rangle_\lambda(\hat{\mathbf{X}}) &= \frac{1}{\pi}\int_{\mathbb{R}^\eta}\sum_{ij} w_{ij}^2\left(\tilde{d}_{ij}-\hat{d}_{ij,u}\right)^2 \exp(\|\mathbf{u}\|_{\mathrm{F}}^2)\, d\mathbf{u},\\ &= \sum_{ij} w_{ij}^2\left(\lambda^2+\tilde{d}_{ij}^2+\hat{d}_{ij}^2-2\lambda\tilde{d}_{ij}\Gamma\left(\frac{3}{2}\right){}_1F_1\left(\frac{3}{2};1;\frac{\hat{d}_{ij}^2}{\lambda^2}\right)\exp\left(\frac{-\hat{d}_{ij}^2}{\lambda^2}\right)\right),\end{aligned} \tag{43}$$

where $\hat{\mathbf{X}} \in \mathbb{R}^{N\times\eta}$ is a matrix whose i-th row-vector is $\hat{\mathbf{x}}_i$, $\hat{d}_{ij,u} \triangleq \|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j + \lambda\mathbf{u}\|_{\mathrm{F}}$ with $\mathbf{u} \in \mathbb{R}^\eta$, $\Gamma(a)$ is the gamma function and ${}_1F_1(a;\, b;\, c)$ is the confluent hypergeometric function Abramowitz & Stegun (1965).

Theorem 3 (Convexity condition) *Let $s(\hat{\mathbf{X}})$ equal to the objective function in equation 7 and $\langle s\rangle_\lambda$ given by 43, then $\langle s\rangle_\lambda$ is convex if*

$$\lambda^* \ge \frac{\sqrt{\pi}\max_{ij}\tilde{d}_{ij}}{2}. \tag{44}$$

Lemma 1 (Minimal $\{\lambda^{(\ell)}\}$ for source localization)
Let $N_{\mathrm{T}} = 1$ and let $\hat{\mathbf{x}}$ denote the target location estimate. Consider an ordered set of ranging measurement $\{\bar{d}_\ell\}$, , such that $\tilde{d}_1 \ge \tilde{d}_2 \ge \ldots \tilde{d}_\ell$, where $L = N_{\mathrm{A}}$. Then the minimal set $\{\lambda^{(\ell)}\}$ is given by

$$\lambda^{(\ell)} = \frac{\sqrt{\pi}\tilde{d}_\ell}{2}, k = 1\ldots L. \tag{45}$$

5.2.4.2 Implementation of the L-GDC

Invoking Theorems 1,2 and 3 the L-GDC algorithm is given by

$$\mathbf{x}^{(\ell)} = \arg\min_{\mathrm{x}\in\mathbb{R}^n}\langle s\rangle_{\lambda^{(\ell)}}(\mathbf{x}),\ 1 \le \ell \le L, \tag{46}$$

with

$$\lambda_0 = \frac{\sqrt{\pi}\max_{ij}\bar{d}_{ij}}{2}. \tag{47}$$

Notice that the ML-transformed objective function involves the hypergeometric function then the numeric evaluation of equation 43 requires care, especially when λ is very small. Numerically stable computations can be achieved using the equivalences Abramowitz & Stegun (1965)

$${}_1F_1\left(\frac{3}{2};1;z\right) = 1 + \sum_{m=1}^{+\infty}\left(z^m \cdot \prod_{k=1}^{m}\frac{(1/2+k)}{k^2}\right), z < 10. \tag{48}$$

$${}_1F_1\left(\frac{3}{2};1;z\right) \approx \frac{z^{-3/2}}{\Gamma(-\frac{1}{2})}\left(\sum_{m=0}^{M-1}\frac{(-z)^{-m}}{m!}\prod_{k=0}^{m-1}\left(\frac{3}{2}+k\right)^2\right) + \frac{e^z z^{1/2}}{\Gamma(\frac{3}{2})}\left(\sum_{p=0}^{P-1}\frac{z^{-p}}{p!}\prod_{k=0}^{p-1}\left(k-\frac{1}{2}\right)^2\right), z \geq 10, \tag{49}$$

where $z \triangleq \frac{\tilde{d}^2}{\lambda^2}$ and M and P are sufficiently large numbers to ensure an accurate approximation (typically $(M,P) \geq 5$).
In order to use a Newton's based optimization method, gradient and Hessian of $\langle s\rangle_{\lambda^{(t)}}(\mathbf{x})$ are required. The gradient is given by

$$\nabla_{\hat{\mathbf{x}}}\langle s\rangle_{\lambda}(\hat{\mathbf{X}}) = \sum_{ij} w_{ij}^2 s'_{ij}(\hat{d}_{ij};\lambda)\nabla_{\hat{\mathbf{x}}}(\hat{d}_{ij}), \tag{50}$$

where the i-th and the j-th $1\times\eta$ blocks of $\nabla_{\hat{\mathbf{x}}}(\hat{d}_{ij})$ are

$$\left[\nabla_{\hat{\mathbf{x}}}(\hat{d}_{ij})\right]_i = \frac{\hat{\mathbf{x}}_j - \hat{\mathbf{x}}_i}{\|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\|_2}, \tag{51}$$

$$\left[\nabla_{\hat{\mathbf{x}}}(\hat{d}_{ij})\right]_j = -\frac{\hat{\mathbf{x}}_j - \hat{\mathbf{x}}_i}{\|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\|_2}, \tag{52}$$

and the function $s'_{ij}(\hat{d}_{ij};\lambda)$ is the first derivative of

$$s_{ij}(\hat{d}_{ij};\lambda) \triangleq \bar{d}_{ij}^2 + \hat{d}_{ij}^2 + \lambda^2 - \lambda\tilde{d}_{ij}\sqrt{\pi}\ {}_1F_1\left(\frac{3}{2},1,\frac{\hat{d}_{ij}^2}{\lambda^2}\right)\exp\left(\frac{-\hat{d}_{ij}^2}{\lambda^2}\right), \tag{53}$$

and is equal to

$$s_{ij'}(\hat{d}_{ij};\lambda) = 2\hat{d}_{ij} + \frac{\hat{d}_{ij}\sqrt{\pi\,d_{ij}}}{\lambda}S_1(\hat{d}_{ij};\lambda), \tag{54}$$

where

$$S_1(\hat{d}_{ij};\lambda) \triangleq \exp\left(\frac{-\hat{d}_{ij}^2}{\lambda^2}\right)\left(2\ {}_1F_1\left(\frac{3}{2};1;\frac{\hat{d}_{ij}^2}{\lambda^2}\right) - 3\ {}_1F_1\left(\frac{5}{2};2;\frac{\hat{d}_{ij}^2}{\lambda^2}\right)\right). \tag{55}$$

The Hessian matrix of $\langle s\rangle_\lambda(\hat{\mathbf{X}})$, denoted by $\nabla^2_{\hat{\mathbf{X}}}\langle s\rangle_\lambda(\hat{\mathbf{X}})$, is computed as

$$\nabla^2_{\hat{\mathbf{X}}}\langle s\rangle_\lambda(\hat{\mathbf{X}}) = \sum_{ij} w_{ij}^2\left(s''_{ij}(\hat{d}_{ij};\lambda)\nabla^{\mathrm{T}}_{\hat{\mathbf{X}}}(\hat{d}_{ij})\nabla_{\hat{\mathbf{X}}}(\hat{d}_{ij}) + s'_{ij}\nabla^2_{\hat{\mathbf{X}}}(\hat{d}_{ij})\right), \tag{56}$$

where $\nabla^2_{\hat{\mathbf{X}}}(\hat{d}_{ij}) \in \mathbb{R}^{N\eta\times N\eta}$ is given by a symmetric block-matrix where the *ii*-th and *ij*-th blocks are

$$\left[\nabla^2_{\hat{\mathbf{X}}}(\hat{d}_{ij})\right]_{ii} = \frac{1}{\hat{d}_{ij}}\left(\mathbf{I} - \left[\nabla_{\hat{\mathbf{X}}}(\hat{d}_{ij})\right]^{\mathrm{T}}_i\left[\nabla^2_{\hat{\mathbf{X}}}(\hat{d}_{ij})\right]_i\right), \tag{57}$$

$$\left[\nabla_{\hat{\mathbf{X}}}(\hat{d}_{ij})\right]_{ij} = -\left[\nabla_{\hat{\mathbf{X}}}(\hat{d}_{ij})\right]_{ii}, \tag{58}$$

and the second derivative of $s_{ij}(\hat{d}_{ij};\lambda)$, denoted by $s''_{ij}(\hat{d}_{ij};\lambda)$, is

$$s''_{ij}(\hat{d}_{ij};\lambda) = 2 + \frac{\sqrt{\pi}\,\bar{d}_{ij}}{\lambda}S_1(\hat{d}_{jj};\lambda) + \frac{\sqrt{\pi}\,\bar{d}_{ij}\hat{d}_{ij}}{\lambda^3}(S_2(\hat{d}_{ij};\lambda) - S_1(\hat{d}_{ij};\lambda)), \tag{59}$$

with

$$S_2(\hat{d}_i;\lambda) \triangleq e^{-\frac{\hat{d}_i^2}{\lambda^2}}\left(3\,{}_1F_1\left(\frac{5}{2};2;\frac{\hat{d}_i^2}{\lambda^2}\right) - \frac{15}{4}\,{}_1F_1\left(\frac{7}{2};3;\frac{\hat{d}_i^2}{\lambda^2}\right)\right). \tag{60}$$

In what follows, we provide an example of source localization problem in $\eta = 1$ dimension. Let $N_A = 2$ and $N_T = 1$. The anchors' and target coordinates are $\mathbf{a}_1 = 0.2$, $\mathbf{a}_2 = 0.5$ and $\mathbf{x} = 0.5$, respectively. We assume no noise, thus $\tilde{d}_i = d_i$. Invoking Theorem 3 we compute $\lambda^{(0)} = 0.3988$ such that $\langle s\rangle_\lambda(\hat{x})$ is convex. Next, we apply the L-GDC technique summarized in equation 46 where, invoking Theorem 2, we choose the set of λ's such that $\lambda^{(0)} = 0.3988$, $\lambda^{(L)} = 0$ and $\lambda^{(\ell)} = \lambda^{(\ell-1)} - 0.05$.

The light-gray lines shown shown in figure 8 indicate the smoothed function computed with the aforementioned set. At each iteration the level of smoothing is decreased and $\langle s\rangle_\lambda(\hat{x})$ approaches more and more $s(\hat{x})$.

The bold lines correspond, instead, to the smoothed function computed for the set of λ's {0.3988, 0.3420, 0.1706,0} using the Lemma 1. In this case, it is shown that $\langle s\rangle_\lambda(\hat{x})$ is recomputed only when a new concave region appears, thus we drastically reduce the computational efficiency of the L-GDC method while preserving optimal performance.

6. Simulation results

In this section, the performance of the non-parametric WLS-ML LT approach considered in this chapter will be evaluated using different optimization algorithms and adopting different weighing strategies. We will use the root-mean-square-error (RMSE) to measure the accuracy of the estimated positions $\hat{\mathbf{X}}$

$$\mathrm{RMSE} \triangleq \sqrt{\frac{1}{RP}\sum_{p=1}^{P}\sum_{r=1}^{R}\|\hat{\mathbf{X}}_{rp} - \mathbf{X}\|_2^2}, \tag{61}$$

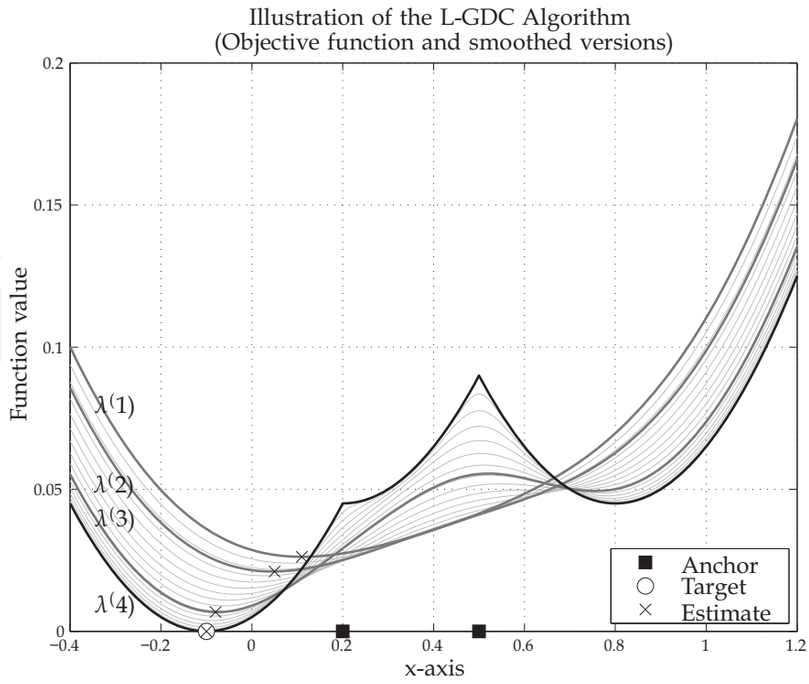


Fig. 8. Illustration of the L-GDC method and the smoothing process. Light-gray lines indicate smoothed versions of the objective functions obtained with a linear decreasing sequence of λ's. Bold lines indicate the smoothed objective with the optimized λ selection criteria.

where R and P are respectively, the number of realizations and networks considered.
For the purpose of comparison, we also benchmark the results to the Cramér-Rao lower bound (CRLB) derived in Jourdan et al. (2006) Patwari et al. (2003) and given by

$$\text{CRLB} \triangleq \text{tr}(\mathbf{F}^{\dagger}), \tag{62}$$

where $\mathbf{F}$ is the Fisher information matrix, that for $\eta = 2$ is equal to

$$\mathbf{F} \triangleq \begin{bmatrix} \mathbf{F}_{xx} & \mathbf{F}_{xy} \\ \mathbf{F}_{xy}^{\mathrm{T}} & \mathbf{F}_{yy} \end{bmatrix}, \tag{63}$$

where

$$\left[\mathbf{F}_{xx}\right]_{jl} = \begin{cases} \sum_{e_j \in E} \dfrac{K_{il}}{\sigma_{il}^2} \dfrac{(x_l - x_i)^2}{d_{il}^2}, & j = l \\ -\dfrac{K_{il}}{\sigma_{il}^2} \dfrac{(x_l - x_j)^2}{d_{jl}^2}, & j \neq l \ \text{ and } \ e_{jl} \in E \end{cases} \tag{64}$$

$$\left[\mathbf{F}_{yy}\right]_{jl} = \begin{cases} \sum_{e_j \in E} \dfrac{K_{il}}{\sigma_{il}^2} \dfrac{(y_l - y_i)^2}{d_{il}^2}, & j = l \\ -\dfrac{K_{il}}{\sigma_{il}^2} \dfrac{(y_l - y_j)^2}{d_{jl}^2}, & j \neq l \ \text{ and } \ e_{jl} \in E \end{cases} \tag{65}$$

$$\left[\mathbf{F}_{xy}\right]_{jl} = \begin{cases} \sum_{e_j \in E} \dfrac{K_{il}}{\sigma_{il}^2} \dfrac{(x_l - x_i)(y_l - y_i)}{d_{il}^2}, & j = l \\ -\dfrac{K_{il}}{\sigma_{il}^2} \dfrac{(x_l - x_j)(y_l - y_j)}{d_{jl}^2}, & j \neq l \text{ and } e_{jl} \in E \end{cases} \tag{66}$$

where e_j indicates the set of links connected to the j-th node.

The first case-of-study is a network with $N_A = 4$ anchors and one target deployed in a square area of size [–10,10] × [–10,10]. The target location is generated as a random variable with uniform distribution within the size of the square while anchors, are located at the locations $\mathbf{x}_1$ = [–10,–10], $\mathbf{x}_2$ = [10,–10], $\mathbf{x}_3$ = [10,10] and $\mathbf{x}_4$ = [–10,10]. We assume that all nodes are connected and the distance of each link is measured K_{ij} times, with $K_{ij} \in$ [2,7]. We use the ranging model given in equation 2 to generate distance measurements, and we consider $\sigma_{ij} \in$ (1e-4,$\sigma_{\max}$).

In figure 9, we show the RMSE obtained with different localization algorithms and unitary weight (unweighted strategy). In this particular study, all algorithms have very similar performance, and the reason is due to the convexity property of the WLS-ML objective function. Indeed, if the target is inside the convex-hull formed by the anchors and the noise is not sufficiently large, then the objective function in typically convex. However, all algorithms do not attain the CRLB because, under the assumption that σ_{ij}'s are all different, the unitary weight is not optimal.

In figure 10 we show the RMSE obtained with the L-GDC algorithm using different weighing strategy, namely, the optimal, the unweighted, the exponential and the dispersion weighing strategy given in equations12, 14, 17, and20, respectively. The results show that the L-GDC algorithm using w_{ij}^* is able to achieve the CRLB, whereas the others stay above.

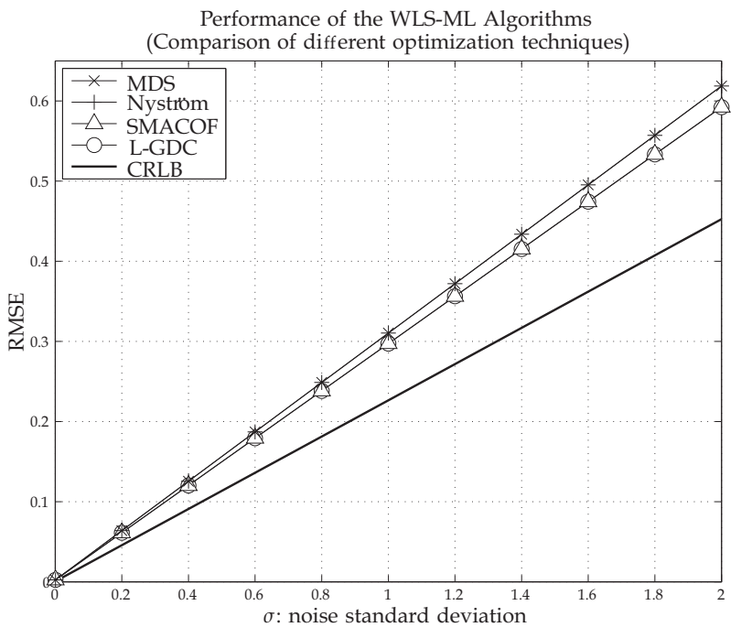


Fig. 9. Comparison of different optimization techniques and using binary weight (unweighted strategy) for a localization problem with $N_A = 4$, $N_T = 1$, $K_{\min} = 2$, $K_{\max} = 7$, $\sigma_{\max} = 1$ and $\sigma_{\min} = 1\text{e-}4$.

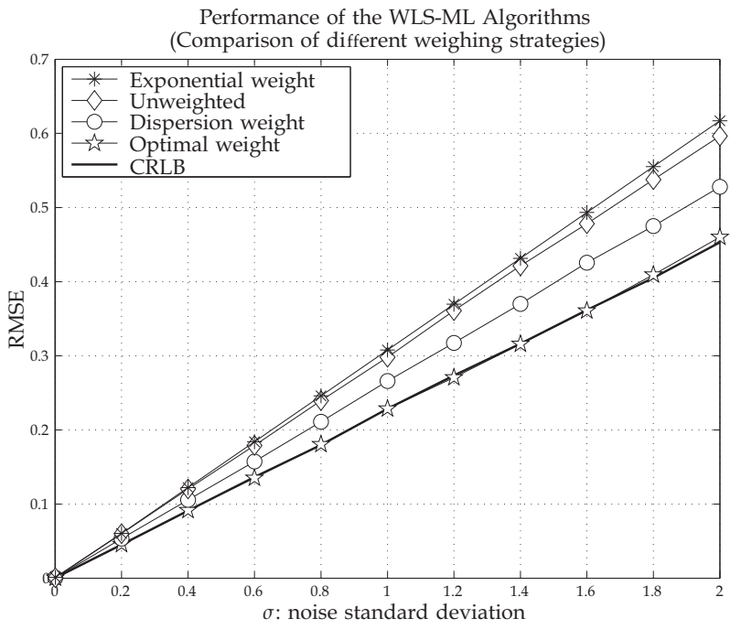


Fig. 10. Comparison of different weighing strategies and using L-GDC optimization method for a localization problem with $N_A = 4$, $N_T = 1$, $K_{min} = 2$, $K_{max} = 7$, $\sigma_{max} = 1$ and $\sigma_{min} = 1e\text{-}4$.

However, to use the optimal weighing strategy we assumed that σ_{ij}'s are known a priori. Therefore, if we reconsider the LT problem under the assumption that the noise statistics are unknown, then the proposed dispersion weight provides the best performance. Indeed, using w_{ij}^{L} we are able to rip ≈ 50% of gain from the unweighted and exponential strategies towards the CRLB.

In the second case-of-study, we consider instead a network with $N_A = 4$ anchors and $N_T = 10$ targets. As before, anchors are located at the corners of a square area while targets are randomly distributed. For this type of simulations, we evaluate the performance of the WLSML algorithms as functions of the *meshness ratio* defined as

$$m \triangleq \frac{(|E| - N + 1)}{(|E_F| - N + 1)}, \tag{67}$$

where E_F indicates the set of links of the fully connected network and $|\cdot|$ indicates the cardinal number of a set Adams & Franzosa (2008)Destino & De Abreu (2009).

This metric is commonly used in algebraic topology and Graph theory to capture, in one number, information on the planarity of a Graph. For example, under the constraint of a connected network, $m = 0$ results from $|E| = N - 1$, which implies that the network is reduced to a tree. In contrast, $m = 1$ results from $|E| = |E_F|$, which implies that the network is not planar, except for the trivial cases of $N \leq 4$. More importantly, the mesheness ratio is an indicator of the connectivity of the network, in a way that is more relevant to its localizability than the simpler connectivity ratio $|E|/|E_F|$.

In figures 11 and 12, the results confirm that the L-GDC is the best optimization technique and, the dispersion weight is the best performing weighing strategy. Similarly to the first case-of-study, also in this case the WLS-ML method based on L-GDC and using the dispersion weights rips about 50% of the error from the alternatives towards the CRLB. Furthermore, from the results shown in figure 11, the L-GDC algorithm is the only one to

maintain an almost constant gap from the CRLB within the entire range of meshness ratio. This let us infer that the L-GDC algorithm finds the global optimum of the WLS-ML function with high probability, while SMACOF of the algebraic methods find sub-optimal solutions.

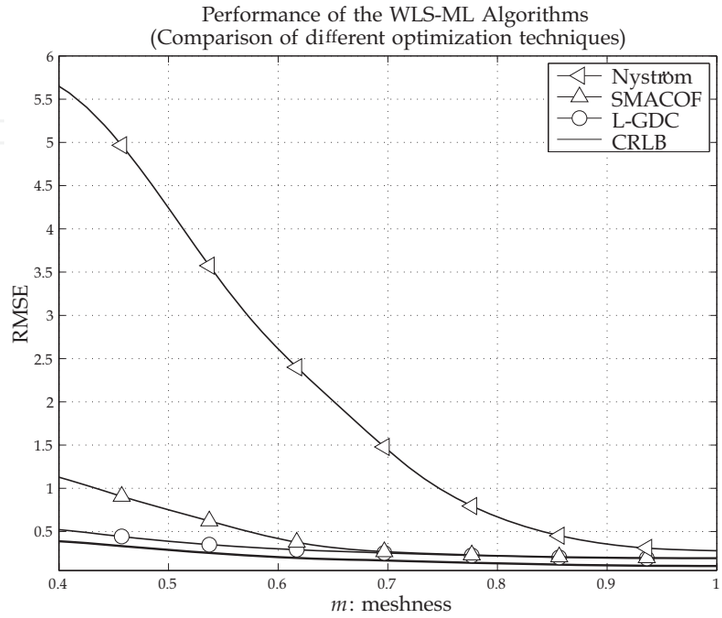


Fig. 11. Comparison of different optimization techniques and using binary weight (unweighted strategy) for a localization problem with $N_A = 4$, $N_T = 10$, $K_{min} = 2$, $K_{max} = 7$, $\sigma_{max} = 1$ and $\sigma_{min} = 1e\text{-}4$.

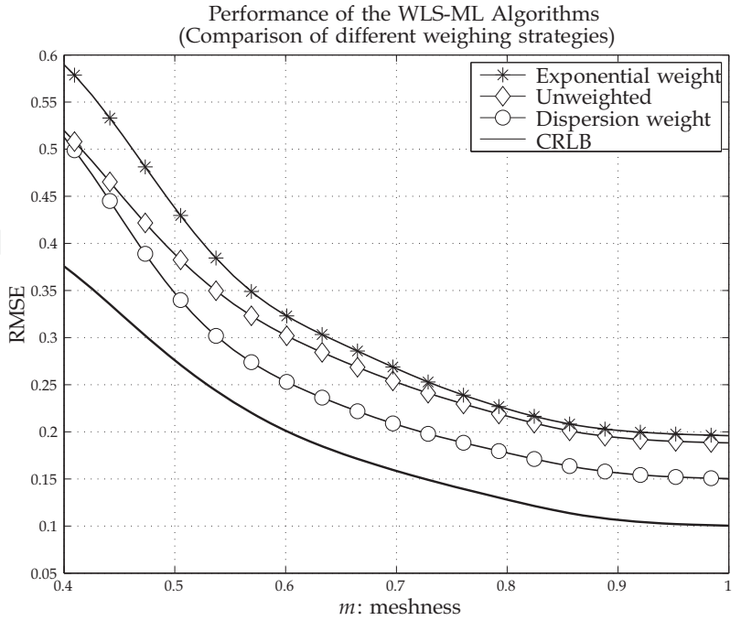


Fig. 12. Comparison of different weighing strategies and using L-GDC optimization method for a localization problem with $N_A = 4$, $N_T = 10$, $K_{min} = 2$, $K_{max} = 7$, $\sigma_{max} = 1$ and $\sigma_{min} = 1e\text{-}4$.

The third and final case-of-study, is the tracking scenario. The network consists of 4 anchor nodes placed at the corner of a square in a $\eta = 2$ dimensional space with 1 targets that moves following an autoregressive model of order 1 within space defined by the anchors. It is assumed full anchor-to-anchor and anchor-to-target connectivity and measurements are perturbed by zero-mean Gaussian noise.
We use the L-GDC optimization method to perform successive re-localization of the target and we employ different weighing strategies. The result shown in figure 14 illustrates the performance of the WLS-ML algorithm as a function of σ considering a velocity $\nu = 1$.
Since the tracking is treated as a mere re-localization, the dynamics only affect the output of the filter block and it is seen from the localization algorithm as an additive noise.
For this reason, the trend of the RMSE is similar to that one obtained in a static scenario. From figure 14 the impact of the velocity on the performance of the WLS-ML algorithm with wavelet-based filter is revealed more clearly. The effect of velocity, indeed, is yet similar to a gaussian noise.
Finally, from both results we observe that the dispersion weight is the best weighing strategy.

7. Conclusions and future work

In this chapter we considered the LT problem in mesh network topologies under LOS conditions. After a general description of the system we focused on a wavelet based filter to smooth the observations and a centralized optimization technique to solve the WLS-ML localization problem. The proposed algorithm was compared with state-of-the-art solutions and it was shown that by combining the wavelet-based filter together with the dispersion weighing strategy and the L-GDC algorithm it is possible to get close to the CRLB.
The work described in this chapter did not address the problem of NLOS channel conditions which needs to be taken into consideration in most of the real life applications. To cope with the biases introduced by NLOS condition two main strategies can be distinguished. In the

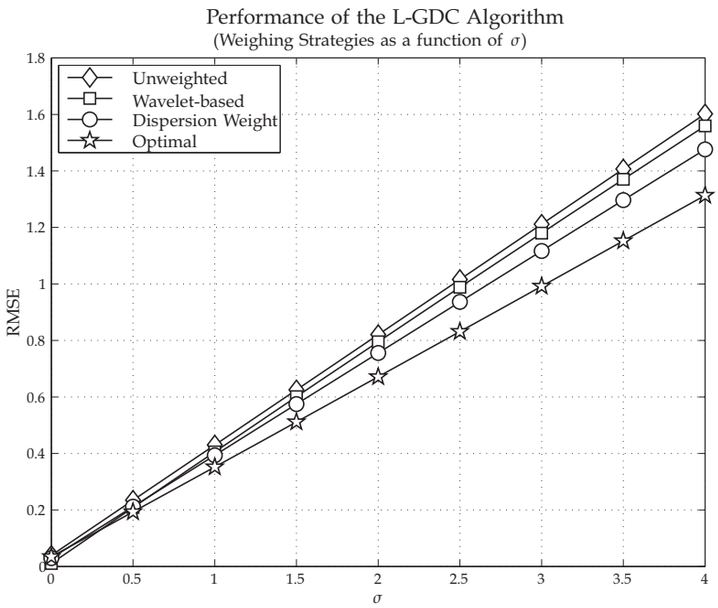


Fig. 13. Performance for the L-GDC algorithm for the different weighing strategies.
Scenario measurements at the 4 anchor nodes subject to normal noise process with standard deviation between 0 and σ.

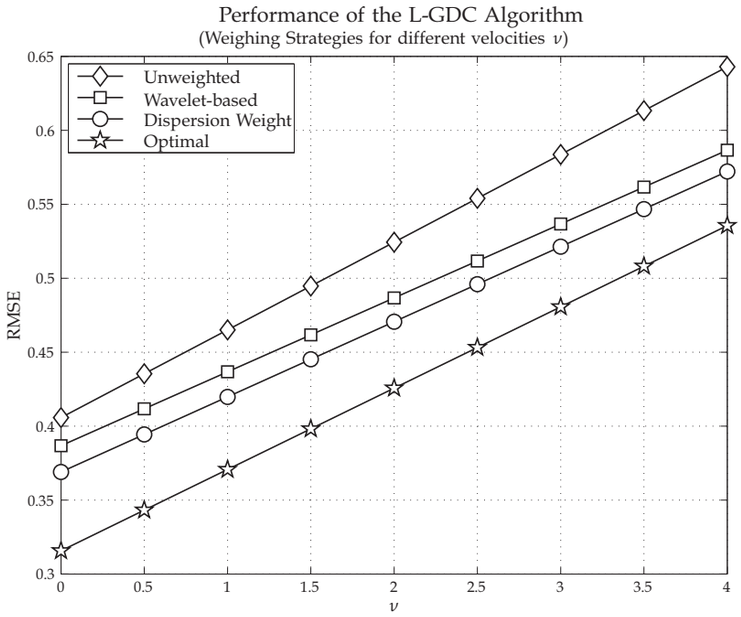


Fig. 14. Performance for the L-GDC algorithm for the different weighing strategies. Scenario measurements at the 4 anchor nodes subject to normal noise process with $\sigma = 2$ and variable target dynamic ν.

first one the biases are treated as additional variables and are directly estimated by the LT algorithm while the second approach aims at discarding the bias introduced by the NLOS condition by applying channel identification and bias compensation algorithms before the LT engine. Concluding, a new method recently proposed by the authors to overcome the NLOS effects is based on an accurate contraction of all the measured distances which has been shown to positively affect the convexity of the objective function and consequently the final location estimates.

8. References

Abramowitz, M. & Stegun, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 10 edn, Dover Publications.

Adams, C. & Franzosa, R. (2008). *Introduction to Topology Pure and Applied*, Pearson Prentice Hall.

Alfakih, A. Y., Wolkowicz, H. & Khandani, A. (1999). Solving euclidean distance matrix completion problems via semidefinite programming, *Journ. on Comp. Opt. and App.* 12(1): 13 – 30.

Beck, A., Stoica, P. & Li, J. (2008). Exact and approximate solutions for source localization problems, *IEEE Trans. Signal Processing* 56(5): 1770–1778.

Biswas, P., Liang, T.-C., Toh, K.-C. & Wang, T.-C. (2006). Semidefinite programming based algorithms for sensor network localization with noisy distance measurements, *ACM Trans. on Sensor Netw. (TOSN)* 2(2): 188–220.

Biswas, P., Liang, T.-C., Toh, K.-C., Wang, T.-C. & Ye, Y. (2006). Semidefinite programming approaches for sensor network localization with noisy distance measurements, *IEEE Trans. Autom. Sci. Eng.* 3: 360–371.

Boyd, S. & Vandenberghe, L. (2004). *Convex Optimization*, Cambridge University Press. C. Fowlkes, S. Belongie, F. C. & Malik, J. (2004). Spectral grouping using the Nystr¨om method, *IEEE Trans. Pattern Anal. Machine Intell.* 26(2).

Cheung, K., So, H., Ma,W.-K. & Chan, Y. (2004). Least squares algorithms for time-of-arrival-based mobile location, *IEEE Trans. on Signal Processing* 52(4): 1121–1130.

Costa, J. A., Patwari, N. & III, A. O. H. (2006). Distributed multidimensional scaling with adaptive weighting for node localization in sensor networks, *ACM J. on Sensor Netw.* 2(1): 39–64.

Cox, T. F. & Cox, M. A. A. (2000). *Multidimensional Scaling*, 2 edn, Chapman & Hall/CRC.

Dattorro, J. (2005). *Convex Optimization and Euclidean Distance Geometry*, Meboo Publishing.

Destino, G. & Abreu, G. (2009). Solving the source localization prolem via global distance continuation, *Proc. IEEE International Conference on Communcations*. IEEE Asilomar Conference on Signals, Systems, and Computers.

Destino, G. & Abreu, G. (2010). On the maximum likelihood formulation of the network localization problem, *(to submit)*.

Destino, G. & De Abreu, G. T. F. (2009). Weighing strategy for network localization under scarce ranging information, *Trans. Wireless. Comm.* 8(7): 3668–3678.

Gibbons, J. (1992). *Nonparametric Statistical Inference*, Marcel Dekker.

Guvenc, I., Gezici, S., Watanabe, F. & Inamura, H. (2008). Enhancements to linear least squares localization through reference selection and ML estimation, *Proc. IEEE Wireless Comm. and Netw. Conf. (WCNC)*, pp. 284–289.

Joon-Yong, L. & Scholtz, R. (2002). Ranging in a dense multipath environment using an UWB radio link., *IEEE J. Sel. Areas Commun.* 20: 1667–1683.

Jourdan, D., Dardari, D. &Win, M. (2006). Position error bound for UWB localization in dense cluttered environments, *Proc. IEEE International Conference on Communcations*, Vol. 8, pp. 3705–3710.

Li, X. & Pahlavan, K. (2004). Super-resolution toa estimation with diversity for indoor geolocation, *IEEE Trans. Wireless Commun.* 3(1): 224–234.

Macagnano, D. & de Abreu, G. T. F. (2008). Tracking multiple dynamic targets in LOS-NLOS condition with multidimensional scaling, *IEEE 5th Workshop on Positioning, Navigation and Communication*.

Mao, G., Fidan, B. & Anderson, B. D. O. (2007). Wireless sensor network localization techniques, *Computer Networks: The Intern. J. of Comp. and Telecomm. Networking* 51(10): 2529–2553.

More, J. &Wu, Z. (1997). Global continuation for distance geometry problems, *SIAM J. Optim.* 7: 814–836.

Nocedal, J. &Wright, S. (2006). *Numerical Optimization*, Springer.

Ouyang, R., Wong, A.-S. & Chin-Tau, L. (2010). Received signal strength-based wireless localization via semidefinite programming: Noncooperative and cooperative schemes, *IEEE Transactions on Vehicular Technology* 59(3): 1307 –1318.

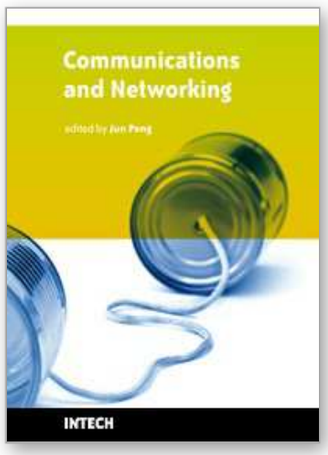
Patwari, N., Dea, R. J. O. & Wang, Y. (2003). Relative location estimation in wireless sensor networks, *IEEE Trans. Signal Processing* 51(8): 2137–2148.

Shang, Y. & Ruml, W. (2004). Improved MDS-based localization, *Proc. 23-rd Ann. Joint Conf. of the IEEE Comp. and Comm. Societies (INFOCOM'04)*, Vol. 4, Hong-Kong, China, pp. 2640 – 2651.

S.Mallat (1998). *A Wavelet Tour of Signal Processing*, second edn, Academic Press.

S.Mallat & S.Zhong (1992). Characterization of signals from multiscale edges, *IEEE Trans. Pattern Anal. Machine Intell.* 14(7): 710–732.

Williams, C. & M.Seeger (2000). Using the Nyström method to speed up kernel machines, *Annual Advances in Neural Information Processing Systems 13* pp. 682–688.



Communications and Networking

Edited by Jun Peng

ISBN 978-953-307-114-5

Hard cover, 434 pages

Publisher Sciyo

Published online 28, September, 2010

Published in print edition September, 2010

This book "Communications and Networking" focuses on the issues at the lowest two layers of communications and networking and provides recent research results on some of these issues. In particular, it first introduces recent research results on many important issues at the physical layer and data link layer of communications and networking and then briefly shows some results on some other important topics such as security and the application of wireless networks. In summary, this book covers a wide range of interesting topics of communications and networking. The introductions, data, and references in this book will help the readers know more abut this topic and help them explore this exciting and fast-evolving field.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Giuseppe Abreu, Giuseppe Destino and Davide Macagnano (2010). Data-Processing and Optimization Methods for Localization-Tracking Systems, Communications and Networking, Jun Peng (Ed.), ISBN: 978-953-307-114-5, InTech, Available from: http://www.intechopen.com/books/communications-and-networking/data-processing-and-optimization-methods-for-localization-tracking-systems

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen