CrossMark

Robust distributed cooperative RSS-based localization for directed graphs in mixed LoS/NLoS environments

Luca Carlino[1], Di Jin[2]* , Michael Muma[2] and Abdelhak M. Zoubir[2]

Abstract

The accurate and low-cost localization of sensors using a wireless sensor network is critically required in a wide range of today's applications. We propose a novel, robust maximum likelihood-type method for distributed cooperative received signal strength-based localization in wireless sensor networks. To cope with mixed LoS/NLoS conditions, we model the measurements using a two-component Gaussian mixture model. The relevant channel parameters, including the reference path loss, the path loss exponent, and the variance of the measurement error, for both LoS and NLoS conditions, are assumed to be unknown deterministic parameters and are adaptively estimated. Unlike existing algorithms, the proposed method naturally takes into account the (possible) asymmetry of links between nodes. The proposed approach has a communication overhead upper-bounded by a quadratic function of the number of nodes and computational complexity scaling linearly with it. The convergence of the proposed method is guaranteed for compatible network graphs, and compatibility can be tested a priori by restating the problem as a graph coloring problem. Simulation results, carried out in comparison to a centralized benchmark algorithm, demonstrate the good overall performance and high robustness in mixed LoS/NLoS environments.

Keywords: Cooperative localization, Received signal strength (RSS), Maximum likelihood estimation, Wireless sensor network (WSN)

1 Introduction

The wide spread of telecommunication systems has led to the pervasiveness of radiofrequency (RF) signals in almost every environment of daily life. Knowledge of the location of mobile devices is required or beneficial in many applications [1], and numerous localization techniques have been proposed over the years [1–4]. Techniques based on the received signal strength (RSS) are the preferred option when low cost, simplicity, and technology obliviousness are the main requirements. In some standards, e.g., IEEE 802.15.4, an RSS indicator (RSSI) is encoded directly into the protocol stack [5]. In addition, RSS is readily available from any radio interface through a simple energy detector and can be modeled by the well-known path loss model [6] regardless of the particular communication

scheme. Based on that, RSS can be exploited to implement "opportunistic" localization for different wireless technologies, e.g., Wi-Fi [7], FM radio [8], or cellular networks [9]. In the context of wireless sensor networks (WSNs), nodes with known positions (*anchors*) can be used to localize the nodes with unknown positions (*agents*). Generally speaking, localization algorithms can be classified according to three important categories.

Centralized vs. distributed. Centralized algorithms, e.g., [10–14], require a data fusion center that carries out the computation after collecting information from the nodes, while distributed algorithms, such as [15, 16], rely on self-localization and the computation is spread throughout the network. Centralized algorithms are likely to provide more accurate estimates, but they suffer from scalability problems, especially for large-scale WSNs, while distributed algorithms have the advantage of being scalable and more robust to node failures [17].

Cooperative vs. non-cooperative. In a non-cooperative algorithm, e.g., [18], each agent receives information only

*Correspondence: djin@spg.tu-darmstadt.de
[2]Signal Processing Group, Department of Electrical Engineering and Information Technology, Technische Universität Darmstadt, Darmstadt, Germany
Full list of author information is available at the end of the article

from anchors. For all agents to obtain sufficient information to perform localization, non-cooperative algorithms necessitate either long range (and high-power) anchor transmission or a high-density of anchors [17]. In cooperative algorithms, such as [19, 20], inter-agent communication removes the need for all agents to be in range of one (or more) anchors [17].

Bayesian vs. non-Bayesian. In non-Bayesian algorithms, e.g., expectation-maximization (EM) [21], and its variant, expectation-conditional maximization (ECM) [22], the unknown positions are treated as deterministic, while in Bayesian algorithms, e.g., nonparametric belief propagation (NBP) [23] and sum-product algorithm over wireless networks (SPAWN) [24, 25] and its variant Sigma-Point SPAWN [26], the unknown positions are assumed to be random variables with a known prior distribution.

Many existing works on localization using RSS measures, such as [27] and [28], are based on the assumption that the classical path loss propagation model is perfectly known, mostly via a calibration process. However, this assumption is impractical for two reasons. Firstly, conducting the calibration process requires intensive human assistance, which may be not affordable, or may even be impossible in some inaccessible areas. Secondly, the channel characteristics vary due to multipath (fading), and non-negligible modifications occur also due to mid- to long-term changes in the environment, leading to non-stationary channel parameters [29]. This implies that the calibration must be performed continuously [29, 30] because otherwise the resulting mismatch between design assumptions and actual operating conditions leads to severe performance degradation. These facts highlight the need for algorithms that *adaptively* estimate the environment and the locations. A further difficulty is due to the existence of non-line-of-sight (NLoS) propagation in practical localization environments. Among various works handling the NLoS effect, a majority of them have treated the NLoS meaures as outliers and tried to neglect or mitigate their effect, including the maximum likelihood (ML)-based approach [31, 32], the weighted least-squares (WLS) estimator [32, 33], the constrained localization techniques [34, 35], robust estimators [36, 37], and the method based on virtual stations [38]. In contrast to these works, several approaches, including [21, 22, 39], have proposed specific probabilistic models for the NLoS measures, therewith exploiting the NLoS measures for the localization purpose. In the light of these considerations, our aim is to develop an RSS-based, cooperative localization framework that works in mixed LoS/NLoS environments, requires no knowledge on parameters of the propagation model, and can be realized in a distributed manner.

The following distinction is made: only algorithms that directly use RSS measures as inputs are considered *RSS-based* in the strict sense, while algorithms such as NBP [23], SPAWN [24], distributed-ECM (D-ECM) [22], and their variants are here called *pseudo-RSS-based*, because they use range estimates as inputs. Generally speaking, among these two options, RSS-based location estimators are preferred for the following reasons. Firstly, inferring the range estimates from the RSS measures usually requires knowledge on the (estimated) propagation model parameters. The assumption of a priori known parameters violates the calibration-free requirement. Secondly, even with perfectly known model parameters, there exists no efficient estimator for estimating the ranges from the RSS measures, as proven in [40]. Thirdly, dropping the idealistic assumption of known channel parameters and using their estimates introduce an irremovable large bias, as demonstrated in [41]. Based on these considerations, pseudo-RSS-based approaches do not meet the requirements in this work. Furthermore, Bayesian approaches, including NBP [23] and SPAWN [24, 25], do not consider mixed LoS/NLoS environments. As one representative RSS-based cooperative localization algorithm, Tomic et al.'s semidefinite-programming (SDP) estimator in [20] requires no knowledge on the propagation model, but it does not apply to and cannot be readily extended to a mixed LoS/NLoS environment. To the best of our knowledge, the existing works on RSS-based calibration-free localization in a mixed LoS/NLoS environment is rather limited. Yin et al. have proposed an EM-based estimator in [21], but only for the single-agent case. In this paper, we consider a multi-agent case and aim to develop a location estimator that is RSS-based, cooperative, and calibration-free and works in a mixed LoS/NLoS environment. To capture the mixed LoS/NLoS propagation conditions, we adopt the mode-dependent propagation model in [21]. The key difference between this work and [21] lies in whether the localization environment is cooperative or not. More precisely, [21] is concerned with the conventional single-agent localization while this work studies cooperative localization in case of multi-agent. Furthermore, we develop a distributed algorithm, where model parameters and positions are updated locally by treating the estimated agents as anchors, inspired by the works in [42–44]. A succinct characterization of the proposed work and its related works is listed in Table 1.

Original contributions: We address the problem of RSS-based cooperative localization in a mixed LoS/NLoS propagation environment, requiring no calibration. To characterize such a mixed LoS/NLoS environment, we assume a mode-dependent propagation model with unknown parameters. We derive and analyze a robust, calibration-free, RSS-based distributed cooperative algorithm, based on the ML framework, which is capable

Table 1 Succinct characterization of the related works and the proposed work

Paper	RSS-based	Cooperative	Calibration-free	LoS/NLoS	Distributed
NBP [23]	No	Yes	N/A	No	Yes
SPAWN [24]	No	Yes	N/A	No	Yes
D-ECM [22]	No	Yes	N/A	Yes	Yes
SDP [20]	Yes	Yes	Yes	No	No
EM [21]	Yes	No	Yes	Yes	N/A
Proposed work	Yes	Yes	Yes	Yes	Yes

N/A not applicable. The terms "RSS-based," "Cooperative," "Calibration-free," "LoS/NLoS," and "Distributed" are defined and discussed in Section 1

of coping with mixed LoS/NLoS conditions. Simulation results, carried out in comparison with a centralized ML algorithm that serves as a benchmark, show that the proposed approach has good overall performance. Moreover, it adaptively estimates the channel parameters, has acceptable communication overhead and computation costs, thus satisfying the major requirements of a practically viable localization algorithm. The convergence analysis of the proposed algorithm is conducted by restating the problem as a graph coloring problem. In particular, we formulate a graph compatibility test and show that for compatible network structures, the convergence is guaranteed. Unlike existing algorithms, the proposed method naturally takes into account the (possible) asymmetry of links between nodes.

The paper is organized as follows. Section 2 formulates the problem and details the algorithms. Section 3 discusses convergence. Section 4 presents the simulations results, while Section 5 concludes the paper. Finally, Appendices A and B contain some analytical derivations which would otherwise burden the reading of the paper.

2 Methods/experimental
2.1 Problem formulation
Consider[1] a directed graph with N_a anchor nodes and N_u agent nodes, for a total of $N = N_a + N_u$ nodes. In a two-dimensional (2D) scenario, we denote the position of node i by $\boldsymbol{x}_i = [x_i \ y_i]^\top \in \mathbb{R}^{2 \times 1}$, where $^\top$ denotes transpose. Between two distinct nodes i and j, the binary variable $o_{j \to i}$ indicates if a measure, onto direction $j \to i$, is observed ($o_{j \to i} = 1$) or not ($o_{j \to i} = 0$). In the case when $i = j$, since a node does not self-measure, we have $o_{j \to i} = 0$. This allows us to define the observation matrix $\mathcal{O} \in \mathbb{B}^{N \times N}$ with elements $o_{i,j} \triangleq o_{i \to j}$ as above. The aforementioned directed graph has connection matrix $\mathcal{O}$. It is important to remark that, for a directed graph, $\mathcal{O}$ is not necessarily symmetric; physically, this models possible

channel anisotropies, miss-detections and, more generally, link failures. Let $m_{j \to i}$ be a binary variable, which denotes if the link $j \to i$ is LoS ($m_{j \to i} = 1$) or NLoS ($m_{j \to i} = 0$). Due to physical reasons, $m_{j \to i} = m_{i \to j}$. We define the LoS/NLoS matrix[2] $\mathbf{L} \in \mathbb{B}^{N \times N}$ of elements $l_{i,j} \triangleq m_{i \to j}$, and we observe that, since $m_{j \to i} = m_{i \to j}$, the matrix is symmetric, i.e., $\mathbf{L}^\top = \mathbf{L}$. We stress that this symmetry is preserved regardless of $\mathcal{O}$, as it derives from physical reasons only. Let $\Gamma(i)$ be the *(open) neighborhood* of node i, i.e., the set of all nodes from which node i receives observables (RSS measures), formally: $\Gamma(i) \triangleq \{j \neq i : o_{j \to i} = 1\}$. We define $\Gamma_a(i)$ as the anchor-neighborhood of node i, i.e., the subset of $\Gamma(i)$ which contains only anchor nodes as neighbors of node i. We also define $\Gamma_u(i)$ as the agent-neighborhood of node i, i.e., the subset of $\Gamma(i)$ which contains only agent nodes as neighbors of node i. In general, $\Gamma(i) = \Gamma_a(i) \cup \Gamma_u(i)$.

2.2 Data model
In the sequel, we will assume that all nodes are stationary and that the observation time-window is sufficiently short in order to neglect correlation in the shadowing terms. In practice, such a model simplification allows for a more analytical treatment of the localization problem and has also been used, for example, in [11, 12]. Following the path loss model and the data models present in the literature [17, 21] and denoting by K the number of samples collected on each link over a predetermined time window, we model the received power at time index k for anchor-agent links as

$$r_{a \to i}^{(m)}(k) = \begin{cases} p_{0_{\text{LOS}}} - 10\alpha_{\text{LOS}} \log_{10} \|\boldsymbol{x}_a - \boldsymbol{x}_i\| + w_{a \to i}(k), \\ \quad \text{if } m_{a \to i} = 1; \\ p_{0_{\text{NLOS}}} - 10\alpha_{\text{NLOS}} \log_{10} \|\mathbf{x}_a - \mathbf{x}_i\| + v_{a \to i}(k), \\ \quad \text{if } m_{a \to i} = 0, \end{cases}$$

(1)

while, for the agent-agent link,

$$r_{u \to i}^{(m)}(k) = \begin{cases} p_{0_{\text{LOS}}} - 10\alpha_{\text{LOS}} \log_{10} \|\boldsymbol{x}_u - \boldsymbol{x}_i\| + w_{u \to i}(k), \\ \quad \text{if } m_{u \to i} = 1; \\ p_{0_{\text{NLOS}}} - 10\alpha_{\text{NLOS}} \log_{10} \|\boldsymbol{x}_u - \boldsymbol{x}_i\| + v_{u \to i}(k), \\ \quad \text{if } m_{u \to i} = 0, \end{cases}$$

(2)

where:

- i, u, with $u \in \Gamma_u(i)$, are the indexes for the unknown nodes;
- $a \in \Gamma_a(i)$ is an index for anchors;
- $k = 1, \dots, K$ is the discrete time index, with K samples for each link;
- $p_{0_{\text{LOS/NLOS}}}$ is the reference power (in dBm) for the LoS or NLoS case;

- $\alpha_{\text{LOS/NLOS}}$ is the path loss exponent for the LoS or NLoS case;
- $\mathbf{x}_a$ is the known position of anchor a;
- $\mathbf{x}_u$ is the unknown position of agent u (similarly for $\boldsymbol{x}_i$);
- $\Gamma_a(i)$, $\Gamma_u(i)$ are the anchor- and agent-neighborhoods of node i, respectively;
- The noise terms $w_{a\to i}(k)$, $v_{a\to i}(k)$, $w_{u\to i}(k)$, and $v_{u\to i}(k)$ are modeled as serially independent and identically distributed (i.i.d.), zero-mean, Gaussian random variables, independent from each other (see below), with variances:
  $\text{Var}[w_{a\to i}(k)] = \text{Var}[w_{u\to i}(k)] = \sigma_{\text{LOS}}^2$,
  $\text{Var}[v_{a\to i}(k)] = \text{Var}[v_{u\to i}(k)] = \sigma_{\text{NLOS}}^2$,
  and $\sigma_{\text{NLOS}}^2 > \sigma_{\text{LOS}}^2 > 0$.

More precisely, letting $\delta_{i,j}$ be Kronecker's delta[3], the independence assumption is formalized by the following equations

$$
\begin{aligned}
\mathbb{E}\left[w_{j_1\to i_1}(k_1)w_{j_2\to i_2}(k_2)\right] &= \sigma_{\text{LOS}}^2\delta_{k_1,k_2}\delta_{i_1,i_2}\delta_{j_1,j_2} \\
\mathbb{E}\left[v_{j_1\to i_1}(k_1)v_{j_2\to i_2}(k_2)\right] &= \sigma_{\text{NLOS}}^2\delta_{k_1,k_2}\delta_{i_1,i_2}\delta_{j_1,j_2} \\
\mathbb{E}\left[w_{j_1\to i_1}(k_1)v_{j_2\to i_2}(k_2)\right] &= 0
\end{aligned}
\tag{3}
$$

for any $k_1, k_2, i_1, i_2, j_1 \in \Gamma(i_1), j_2 \in \Gamma(i_2)$. The previous equations imply that two different links are always independent, regardless of the considered time instant. In this paper, we call this property *link independence*. If only one link is considered, i.e., $j_2 = j_1$ and $i_2 = i_1$, then independence is preserved by choosing different time instants, implying that the sequence $\{w_{j\to i}\}_k \triangleq \{w_{j\to i}(1), w_{j\to i}(2), \dots\}$ is white. The same reasoning applies to the (similarly defined) sequence $\{v_{j\to i}\}_k$. As a matter of notation, we denote the unknown positions (indexing the agents before the anchors) by $\boldsymbol{x} \triangleq \left[\boldsymbol{x}_1^\top \cdots \boldsymbol{x}_{N_u}^\top\right]^\top \in \mathbb{R}^{2N_u\times 1}$ and we define $\boldsymbol{\eta}$ as the collection of all channel parameters, i.e., $\boldsymbol{\eta} \triangleq \left[\boldsymbol{\eta}_{\text{LOS}}^\top \ \boldsymbol{\eta}_{\text{NLOS}}^\top\right]^\top$, with $\boldsymbol{\eta}_{\text{LOS}} \triangleq \left[p_{0_{\text{LOS}}} \ \alpha_{\text{LOS}} \ \sigma_{\text{LOS}}^2\right]^\top \in \mathbb{R}^{3\times 1}$, $\boldsymbol{\eta}_{\text{NLOS}} \triangleq \left[p_{0_{\text{NLOS}}} \ \alpha_{\text{NLOS}} \ \sigma_{\text{NLOS}}^2\right]^\top \in \mathbb{R}^{3\times 1}$.

It is important to stress that, in a more realistic scenario, channel parameters may vary from link to link and also across time. However, such a generalization would produce an under-determined system of equations, thus giving up uniqueness of the solution and, more generally, analytical tractability of the problem. For the purposes of this paper, the observation model above is sufficiently general to solve the localization task while retaining analytical tractability.

2.3 Time-averaged RSS measures
Motivated by a result given in Appendix A, we consider the time-averaged RSS measures, defined as

$$
\bar{r}_{j\to i} \triangleq \frac{1}{K}\sum_{k=1}^{K} r_{j\to i}^{(m)}(k), \quad j \in \Gamma(i)
\tag{4}
$$

as our new observables[4]. While it would have been preferable to work with the original data from a theoretical standpoint, several considerations lead to the preference of time-averaged data, most notably: (1) comparison with other algorithms present in the literature, where the data model assumes only one sample per link, i.e., $K = 1$, which is simply a special case in this paper; (2) reduced computational complexity in the subsequent algorithms; (3) if the RSS measures onto a given link needs to be communicated between two nodes, the communication cost is notably reduced, since only one scalar, instead of K samples, needs to be communicated; (4) formal simplicity of the subsequent equations.

Moreover, from Appendix A, it follows that, assuming known $\boldsymbol{L}$, the ML estimators of the unknown positions based upon an original data or a time-averaged data are actually the same. To see this, it suffices to choose $\boldsymbol{\theta} = (\boldsymbol{x}, p_{0_{\text{LOS}}}, p_{0_{\text{NLOS}}}, \alpha_{\text{LOS}}, \alpha_{\text{NLOS}})$ and

$$
s_{j\to i}(\boldsymbol{\theta}) =
\begin{cases}
p_{0_{\text{LOS}}} - 10\alpha_{\text{LOS}}\log_{10}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|, \\
\qquad \text{if } m_{j\to i} = 1; \\
p_{0_{\text{NLOS}}} - 10\alpha_{\text{NLOS}}\log_{10}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|, \\
\qquad \text{if } m_{j\to i} = 0
\end{cases}
\tag{5}
$$

for $j \in \Gamma(i)$ and splitting the additive noise term as required. For a fixed link, only one of two cases (LoS or NLoS) is verified, thus applying (34) of Appendix A yields

$$
\begin{aligned}
\arg\max_{\boldsymbol{\theta}} \ & p\left(r_{j\to i}^{(m)}(1), \dots, r_{j\to i}^{(m)}(K); \boldsymbol{\theta}, \sigma_{j\to i}^2\right) \\
&= \arg\max_{\boldsymbol{\theta}} p\left(\bar{r}_{j\to i}; \boldsymbol{\theta}, \sigma_{j\to i}^2\right)
\end{aligned}
\tag{6}
$$

where $\sigma_{j\to i}^2$ is either σ_{LOS}^2 or σ_{NLOS}^2 and the general result follows from link independence.

We define $\mathcal{R}_i$ as the set of all RSS measures that node i receives from anchor neighbors, i.e., $\mathcal{R}_i \triangleq \left\{r_{a\to i}^{(m)}(1), \dots, r_{a\to i}^{(m)}(K) : a \in \Gamma_a(i)\right\}$, $\mathcal{Z}_i$ as the set of all RSS measures that node i receives from agent neighbors, i.e., $\mathcal{Z}_i \triangleq \left\{r_{j\to i}^{(m)}(1), \dots, r_{j\to i}^{(m)}(K) : j \in \Gamma_u(i)\right\}$, and $\mathcal{Y}_i$ as the set of all RSS measures locally available to node i, i.e., $\mathcal{Y}_i \triangleq \mathcal{R}_i \cup \mathcal{Z}_i$. Analogously, for time-averaged measures, we define $\bar{\mathcal{R}}_i \triangleq \{\bar{r}_{a\to i} : a \in \Gamma_a(i)\}$, $\bar{\mathcal{Z}}_i \triangleq \{\bar{r}_{j\to i} : j \in \Gamma_u(i)\}$, and $\bar{\mathcal{Y}}_i = \bar{\mathcal{R}}_i \cup \bar{\mathcal{Z}}_i$. Finally, we define

$$
\Upsilon \triangleq \bigcup_{i=1}^{N_u} \mathcal{Y}_i
\tag{7}
$$

which represents the information available to the whole network.

2.4 Single-agent robust maximum likelihood (ML)

We first consider the single-agent case, which we will later use as a building block in the multi-agent case. The key idea is that instead of separately treating the LoS and NLoS cases, e.g., by hypothesis testing, we resort to a two-component Gaussian mixture model for the time-averaged RSS measures. More precisely, we assume that the probability density function (pdf), $p(\cdot)$, of the time-averaged RSS measures, for anchor-agent links, is given by

$$p(\bar{r}_{a\to i})$$

$$= \frac{\lambda_i}{\sqrt{2\pi \frac{\sigma_{\text{LOS}}^2}{K}}} e^{-\frac{K}{2\sigma_{\text{LOS}}^2}\left(\bar{r}_{a\to i} - p_{0_{\text{LOS}}} + 10\alpha_{\text{LOS}}\log_{10}\|x_a - x_i\|\right)^2}$$

$$+ \frac{1-\lambda_i}{\sqrt{2\pi \frac{\sigma_{\text{NLOS}}^2}{K}}} e^{-\frac{K}{2\sigma_{\text{NLOS}}^2}\left(\bar{r}_{a\to i} - p_{0_{\text{NLOS}}} + 10\alpha_{\text{NLOS}}\log_{10}\|x_a - x_i\|\right)^2}$$

$$(8)$$

and, for agent-agent links,

$$p(\bar{r}_{u\to i})$$

$$= \frac{\zeta_i}{\sqrt{2\pi \frac{\sigma_{\text{LOS}}^2}{K}}} e^{-\frac{K}{2\sigma_{\text{LOS}}^2}\left(\bar{r}_{u\to i} - p_{0_{\text{LOS}}} + 10\alpha_{\text{LOS}}\log_{10}\|x_u - x_i\|\right)^2}$$

$$+ \frac{1-\zeta_i}{\sqrt{2\pi \frac{\sigma_{\text{NLOS}}^2}{K}}} e^{-\frac{K}{2\sigma_{\text{NLOS}}^2}\left(\bar{r}_{u\to i} - p_{0_{\text{NLOS}}} + 10\alpha_{\text{NLOS}}\log_{10}\|x_u - x_i\|\right)^2}$$

$$(9)$$

where:

- $\lambda_i \in (0,1)$ is the mixing coefficient for anchor-agent links of node i;
- $\zeta_i \in (0,1)$ is the mixing coefficient for agent-agent links of node i.

Empirically, we can intuitively interpret λ_i as the fraction of anchor-agent links in LoS (for node i), while ζ_i as the fraction of agent-agent links in LoS (for node i). As in [21], the Markov chain induced by our model is regular and time-homogeneous. From this, it follows that the Markov chain will converge to a two-component Gaussian mixture, giving a theoretical justification to the proposed approach.

Assume that there is a single agent, say node i, with a minimum of three anchors[5] in its neighborhood ($|\Gamma_a(i)| \geq 3$), in a mixed LoS/NLoS scenario. Our goal is to obtain the maximum likelihood estimator (MLE) of the position of node i. Let $\bar{r}_i = \begin{bmatrix} \bar{r}_{1\to i} & \cdots & \bar{r}_{|\Gamma_a(i)|\to i} \end{bmatrix}^\top \in \mathbb{R}^{|\Gamma_a(i)|\times 1}$ be the collection of all the time-averaged RSS measures available to node i. Using the previous assumptions and

the independency between the links, the joint likelihood function[6] $p(\bar{r}_i; \boldsymbol{\theta})$ is given by

$$p(\bar{r}_i; \boldsymbol{\theta}) = \prod_{a\in\Gamma_a(i)} p(\bar{r}_{a\to i}; \boldsymbol{\theta}) \tag{10}$$

where $\boldsymbol{\theta} = (x_i, \lambda_i, \boldsymbol{\eta})$. Thus, denoting with $L(\boldsymbol{\theta}; \bar{r}_i)$ the log-likelihood, we have

$$L(\boldsymbol{\theta}; \bar{r}_i) = \sum_{a\in\Gamma_a(i)} \ln p(\bar{r}_{a\to i}) \tag{11}$$

The MLE of $\boldsymbol{\theta}$ is given by

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} L(\boldsymbol{\theta}; \bar{r}_i) \tag{12}$$

where the maximization is subject to several constraints: $\lambda_i \in (0,1), \alpha_{\text{LOS}} > 0, \alpha_{\text{NLOS}} > 0, \sigma_{\text{LOS}}^2 > 0$, and $\sigma_{\text{NLOS}}^2 > 0$. In general, the previous maximization admits no closed-form solution, so we must resort to numerical procedures.

2.5 Multi-agent robust ML-based scheme

In principle, our goal would be to have a ML estimate of all the N_u unknown positions, denoted by x. Let $\boldsymbol{\lambda} \triangleq \begin{bmatrix} \lambda_1 & \cdots & \lambda_{N_u} \end{bmatrix}^\top, \boldsymbol{\zeta} \triangleq \begin{bmatrix} \zeta_1 & \cdots & \zeta_{N_u} \end{bmatrix}^\top$ be the collections of the mixing coefficients. Defining $\boldsymbol{\theta} = (x, \boldsymbol{\lambda}, \boldsymbol{\zeta}, \boldsymbol{\eta})$, the ML joint estimator

$$\hat{\boldsymbol{\theta}}_{ML} = \arg\max_{\boldsymbol{\theta}} p(\Upsilon; \boldsymbol{\theta}) \tag{13}$$

is, in general, computationally unfeasible and naturally centralized. In order to obtain a practical algorithm, we now resort to a sub-optimal but computationally feasible and distributed approach. The intuition is as follows. Assume, for a moment, that a specific node i knows $\boldsymbol{\eta}, \boldsymbol{\lambda}, \boldsymbol{\zeta}$ and also all the true positions of its neighbors (which we denote by $\mathcal{X}_i$). Then, the ML joint estimation problem is notably reduced, in fact,

$$\hat{x}_{i_{ML}} = \arg\max_{x_i} p(\Upsilon; x_i) \tag{14}$$

We now make the sub-optimal approximation of avoiding non-local information in order to obtain a distributed algorithm, thus resorting to

$$\hat{x}_i = \arg\max_{x_i} p\left(\bar{\mathcal{Y}}_i; x_i, \mathcal{X}_i, \lambda_i, \zeta_i, \boldsymbol{\eta}\right) \tag{15}$$

where we made explicit the functional dependence on all the other parameters (which, for now, are assumed known). Due to the i.i.d. hypothesis, the "local" likelihood function has the form

$$p\left(\bar{\mathcal{Y}}_i; x_i, \mathcal{X}_i, \lambda_i, \zeta_i, \boldsymbol{\eta}\right) = \prod_{a\in\Gamma_a(i)} p(\bar{r}_{a\to i}; x_i, \lambda_i, \boldsymbol{\eta})$$

$$\times \prod_{j\in\Gamma_u(i)} p(\bar{r}_{j\to i}; x_i, \mathcal{X}_i, \zeta_i, \boldsymbol{\eta}) \tag{16}$$

where the marginal likelihoods are Gaussian-mixtures and we underline the (formal and conceptual) separation between anchor-agent links and agent-agent links. By taking the natural logarithm, we have

$$\ln p\left(\bar{\mathcal{Y}}_i; \boldsymbol{x}_i, \mathcal{X}_i, \lambda_i, \zeta_i, \boldsymbol{\eta}\right) = \sum_{a \in \Gamma_a(i)} \ln p(\bar{r}_{a \to i}; \boldsymbol{x}_i, \lambda_i, \boldsymbol{\eta})$$
$$+ \sum_{j \in \Gamma_u(i)} \ln p(\bar{r}_{j \to i}; \boldsymbol{x}_i, \mathcal{X}_i, \zeta_i, \boldsymbol{\eta})$$
$$(17)$$

The maximization problem in (15) then reads

$$\hat{\boldsymbol{x}}_i = \arg \max_{\boldsymbol{x}_i} \left\{ \sum_{a \in \Gamma_a(i)} \ln p(\bar{r}_{a \to i}; \boldsymbol{x}_i, \lambda_i, \boldsymbol{\eta}) \right.$$
$$\left. + \sum_{j \in \Gamma_u(i)} \ln p(\bar{r}_{j \to i}; \boldsymbol{x}_i, \mathcal{X}_i, \zeta_i, \boldsymbol{\eta}) \right\}. \quad (18)$$

We can now relax the initial assumptions: instead of assuming known neighbors positions $\mathcal{X}_i$, we will substitute them with their estimates, $\hat{\mathcal{X}}_i$. Moreover, since the robust ML-based self-calibration can be done without knowing the channel parameters $\boldsymbol{\eta}$, we also maximize over them. Lastly, we maximize with respect to the mixing coefficients λ_i, ζ_i. Thus, our final approach is

$$\left(\hat{\boldsymbol{x}}_i, \hat{\lambda}_i, \hat{\zeta}_i, \hat{\boldsymbol{\eta}}\right) = \arg \max_{\boldsymbol{x}_i, \lambda_i, \zeta_i, \boldsymbol{\eta}} \left\{ \sum_{a \in \Gamma_a(i)} \ln p(\bar{r}_{a \to i}; \boldsymbol{x}_i, \lambda_i, \boldsymbol{\eta}) \right.$$
$$\left. + \sum_{j \in \hat{\Gamma}_u(i)} \ln p\left(\bar{r}_{j \to i}; \boldsymbol{x}_i, \hat{\mathcal{X}}_i, \zeta_i, \boldsymbol{\eta}\right) \right\}$$
$$(19)$$

where $\hat{\Gamma}_u(i)$ is the set of all agent neighbors of node i for which estimated positions exist. We can iteratively construct (and update) the set $\hat{\Gamma}_u(i)$, in order to obtain a fully distributed algorithm, as summarized in Algorithm 1.

A few remarks are now in order. First, this algorithm imposes some restrictions on the arbitrariness of the network topology, since the information spreads starting from the agents which were able to self-localize during initialization; in practice, this requires the network to be sufficiently connected. Second, convergence of the algorithm is actually a matter of compatibility: if the network is sufficiently connected (compatible), convergence is guaranteed. Given a directed graph, compatibility can be tested a priori and necessary and sufficient conditions can be found (see Section 4). Third, unlike many algorithms present in the literature, symmetrical links are not necessary, nor do we resort to symmetrization (like NBP): this algorithm naturally takes into account the (possible) asymmetrical links of directed graphs.

Algorithm 1 Robust distributed maximum likelihood (RDML)

Initialization:
Self-Localization: Every node i with $|\Gamma_a(i)| \geq 3$ self-localizes by the robust procedure (12);
Broadcast: Every node which self-localized broadcasts its position estimate $\hat{\boldsymbol{x}}_i$;
Iterative scheme: Start with $n \leftarrow 1$;
Update position: Every node with $|\Gamma_a(i)| + |\hat{\Gamma}_u^{(n)}(i)| \geq 3$ estimates its own position by solving (19);
Broadcast: The updated position $\hat{\boldsymbol{x}}_i^{(n)}$ is broadcasted;
Update Neighbors: Every node j for which the new estimated positions $\hat{\boldsymbol{x}}_i^{(n)}$ are neighbors updates its own $\hat{\Gamma}_u^{(n)}(j)$ with the new positions;
Repeat: Set $n \leftarrow n + 1$ and repeat the previous steps until stop condition is met.
Stop condition:
$|\hat{\Gamma}_u^{(n)}(i)| = |\Gamma_u(i)|$ for every node i.

2.6 Distributed maximum likelihood (DML)

As a natural competitor of the proposed RDML algorithm, we derive here the distributed maximum likelihood (DML) algorithm, which assumes that all links are of the same type. As its name suggests, this is the non-robust version of the previously derived RDML. As usual, we start with the single-agent case as a building block for the multi-agent case. Using the assumption that all links are the same and the i.i.d. hypothesis, the joint pdf of the time-averaged RSS measures, received by agent i, is given by

$$p\left(\bar{\boldsymbol{r}}_i; \boldsymbol{x}_i, p_0, \alpha, \sigma^2\right) = \frac{1}{\left(2\pi\sigma^2\right)^{|\Gamma_a(i)|/2}}$$
$$\times e^{-\frac{1}{2\sigma^2} \sum_{a \in \Gamma_a(i)} (\bar{r}_{a \to i} - p_0 + 10\alpha \log_{10} \|\boldsymbol{x}_i - \boldsymbol{x}_a\|)^2}$$
$$(20)$$

We can now proceed by estimating, with the ML criterion, first p_0 as a function of the remaining parameters, followed by α as a function of $\boldsymbol{x}_i$ and finally $\boldsymbol{x}_i$. We have

$$\hat{p}_0(\alpha, \boldsymbol{x}_i) =$$
$$\arg \min_{p_0} \sum_{a \in \Gamma_a(i)} \left(\bar{r}_{a \to i} - p_0 + 10\alpha \log_{10} \|\boldsymbol{x}_i - \boldsymbol{x}_a\|\right)^2. \quad (21)$$

Defining $s_{a,i} \triangleq 10 \log_{10} \|\boldsymbol{x}_i - \boldsymbol{x}_a\|$ as the log-distance, $\boldsymbol{s}_i \triangleq \left[s_{1,i}\ s_{2,i}\ \cdots\ s_{|\Gamma_a(i)|,i}\right]^\top \in \mathbb{R}^{|\Gamma_a(i)| \times 1}$ the column-vector collecting them and $\mathbf{1}_n = [1 \cdots 1]^\top \in \mathbb{R}^{n \times 1}$ an all-ones vector of dimension n, the previous equation can be written as

$$\hat{p}_0(\alpha, \boldsymbol{x}_i) = \arg \min_{p_0} \|\bar{\boldsymbol{r}}_i + \alpha \boldsymbol{s}_i - p_0 \mathbf{1}_{|\Gamma_a(i)|}\|^2 \quad (22)$$

which is a least-squares (LS) problem and its solution is

$$\hat{p}_0(\alpha, \boldsymbol{x}_i) = \frac{1}{|\Gamma_a(i)|} \sum_{a \in \Gamma_a(i)} (\bar{r}_{a \to i} + 10\alpha \log_{10} \|\boldsymbol{x}_i - \boldsymbol{x}_a\|) \tag{23}$$

By using this expression, the problem of estimating α as a function of $\boldsymbol{x}_i$ is

$$\hat{\alpha}(\boldsymbol{x}_i) = \arg\min_{\alpha} \left\| \boldsymbol{P}^{\perp}_{\boldsymbol{1}_{|\Gamma_a(i)|}} (\bar{\boldsymbol{r}}_i + \alpha \boldsymbol{s}_i) \right\|^2 \tag{24}$$

where, given a full-rank matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, with $m \geq n$, $\boldsymbol{P}^{\perp}_{\boldsymbol{A}}$ is the orthogonal projection matrix onto the orthogonal complement of the space spanned by the columns of $\boldsymbol{A}$. It can be computed via $\boldsymbol{P}^{\perp}_{\boldsymbol{A}} = \boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{A}}$, where $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A} \left(\boldsymbol{A}^{\top}\boldsymbol{A}\right)^{-1} \boldsymbol{A}^{\top}$ is an orthogonal projection matrix and $\boldsymbol{I}_m$ is the identity matrix of order m. The solution to problem (24) is given by

$$\hat{\alpha}(\boldsymbol{x}_i) = -\left(\tilde{\boldsymbol{s}}_i^{\top} \tilde{\boldsymbol{s}}_i\right)^{-1} \tilde{\boldsymbol{s}}_i^{\top} \tilde{\boldsymbol{r}}_i \tag{25}$$

where $\tilde{\boldsymbol{r}}_i = \boldsymbol{P}^{\perp}_{\boldsymbol{1}_{|\Gamma_a(i)|}} \bar{\boldsymbol{r}}_i$ and $\tilde{\boldsymbol{s}}_i = \boldsymbol{P}^{\perp}_{\boldsymbol{1}_{|\Gamma_a(i)|}} \boldsymbol{s}_i$. By using the previous expression, we can finally write

$$\hat{\boldsymbol{x}}_i = \arg\min_{\boldsymbol{x}_i} \left\| \boldsymbol{P}^{\perp}_{\tilde{\boldsymbol{s}}_i} \tilde{\boldsymbol{r}}_i \right\|^2 \tag{26}$$

which, in general, does not admit a closed-form solution, but can be solved numerically. After obtaining $\hat{\boldsymbol{x}}_i$, node i can estimate p_0 and α using (23) and (25).

The multi-agent case follows an almost identical reasoning of the RDML. Approximating the true (centralized) MLE by avoiding non-local information and assuming to already have an initial estimate of p_0 and α, it is possible to arrive at

$$\hat{\boldsymbol{x}}_i = \arg\min_{\boldsymbol{x}_i} \left\{ \sum_{a \in \Gamma_a(i)} \left(\bar{r}_{a \to i} - p_0 + 10\alpha \log_{10} \|\boldsymbol{x}_i - \boldsymbol{x}_a\| \right)^2 \right.$$
$$\left. + \sum_{j \in \hat{\Gamma}_u(i)} \left(\bar{r}_{j \to i} - p_0 + 10\alpha \log_{10} \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \right)^2 \right\} \tag{27}$$

where (again) an initialization phase is required and the set of estimated agents-neighbors $\hat{\Gamma}_u(i)$ is iteratively updated. The key difference with RDML is that, due to the assumption of the links being all of the same type, the estimates of p_0 and α are broadcasted and a common consensus is reached by averaging. This increases the communication overhead, but lowers the computational complexity, operating a trade-off. The DML algorithm is summarized in Algorithm 2.

Similar remarks as for the RDML can be made for the DML. Again, the network's topology cannot be completely arbitrary, as the information must spread throughout the network starting from the agents which self-localized,

---

Algorithm 2 Distributed maximum likelihood (DML)

Initialization:

Self-Localization: Every node i with $|\Gamma_a(i)| \geq 3$ self-localizes by solving (26) and estimates α via (25) and p_0 via (23), thus obtaining $\left(\hat{p}_0^{(i)}, \hat{\alpha}^{(i)}, \hat{\boldsymbol{x}}_i\right)$;

Broadcast: Every node which self-localized broadcasts its local estimates $\left(\hat{p}_0^{(i)}, \hat{\alpha}^{(i)}, \hat{\boldsymbol{x}}_i\right)$;

Consensus: All the nodes agree on global values of $(\hat{p}_0, \hat{\alpha})$ by averaging all the local estimates $\left(\hat{p}_0^{(i)}, \hat{\alpha}^{(i)}\right)$;

Iterative scheme: Start with $n \leftarrow 1$;

Update position: Every node i with $\left|\Gamma_a(i)\right| + |\hat{\Gamma}_u^{(n)}(i)| \geq 3$ estimates its own position by solving (27);

Broadcast: The updated position $\hat{\boldsymbol{x}}_i^{(n)}$ is broadcasted;

Update Neighbors: Every node j for which the new estimated positions $\hat{\boldsymbol{x}}_i^{(n)}$ are neighbors updates its own $\hat{\Gamma}_u^{(n)}(j)$ with the new positions;

Repeat: Set $n \leftarrow n + 1$ and repeat the previous steps until stop condition is met.

Stop condition:

$\left|\hat{\Gamma}_u^{(n)}(i)\right| = |\Gamma_u(i)|$ for every node i.

---

implying that the graph must be sufficiently connected. Necessary and sufficient conditions to answer the compatibility question are the same as RDML. Secondly, the (strong) hypothesis behind the DML derivation (i.e., all links of the same type) allows for a more analytical derivation, up to position estimation, which is a nonlinear least-squares problem. However, it is also its weakness since, as will be shown later, it is not a good choice for mixed LoS/NLoS scenarios.

2.7 Centralized MLE with known nuisance parameters (C-MLE)

The centralized MLE of $\boldsymbol{x}$ with known nuisance parameters, i.e., assuming known $\boldsymbol{L}$ and $\boldsymbol{\eta}$, is chosen here as a benchmark for both RDML and DML. In the following, this algorithm will be denoted by C-MLE. Its derivation is simple (see Appendix B) and results in

$$\hat{\boldsymbol{x}}_{ML}$$

$$= \arg\min_{\boldsymbol{x}} \sum_{i=1}^{N_u} \sum_{j \in \Gamma(i)} \frac{1}{\sigma_{j \to i}^2} \left(\bar{r}_{j \to i} - p_{0_{j \to i}} + 10\alpha_{j \to i} \log_{10} \|\boldsymbol{x}_j - \boldsymbol{x}_i\| \right)^2 \tag{28}$$

where $\left(p_{0_{j \to i}}, \alpha_{j \to i}, \sigma_{j \to i}^2\right)$ are either LoS or NLoS depending on the considered link. It is important to observe that, if all links are of the same type, the dependence from $\sigma_{j \to i}^2$ in (28) disappears. From standard ML theory [45], C-MLE is asymptotically ($K \to +\infty$) optimal. The optimization problem (28) is computationally challenging, as

it requires a minimization in a $2N_u$-dimensional space, but still feasible for small values of N_u.

3 Convergence analysis

The convergence test of our proposed algorithm (and also of DML) can be restated as a graph coloring problem: if all the graph can be colored, then it is compatible and convergence is guaranteed. As it is common in the literature on graph theory, let $G = (V, E)$ be a directed graph, with V denoting the set of *nodes* and E the set of *directed edges*. The set of nodes is such that $V = V_a \cup V_u$, where V_a is the (non-empty) set of anchor nodes and V_u is the (non-empty) set of agent nodes.

Definition 1 (RDML-initializable) *A directed graph G is said to be RDML-initializable if and only if there exists at least one agent node, say x, such that $|\Gamma_a(x)| \geq 3$.*

As can be easily checked, the previous statement is a necessary condition: if a graph is not RDML-initializable, then it is incompatible with RDML. To give a necessary and sufficient condition, we introduce the notion of "color." A node can be either black or white; all anchors are black and all agent nodes start as white, but may become black if some condition is satisfied. The RDML can be rewritten as a graph coloring problem. In order to do this, we define the set $\hat{\Gamma}_u(i) \triangleq \{j \in \Gamma_u(i) : \text{agent } j \text{ is black}\}$, i.e., the subset of agent neighbors of node i which contains only black agent nodes. In general, $\hat{\Gamma}_u(i) \subseteq \Gamma_u(i)$. We also define the set $B_u \triangleq \{i \in V_u : \text{agent } i \text{ is black}\}$, i.e., the set of black agents. In general, $B_u \subseteq V_u$. Given a graph G, we can perform a preliminary test by running the following RDML-coloring algorithm (a better test will be derived later):

- **Initialization** $(k = 0)$

    1 All anchors are colored black and all agents white;
    2 Every agent i with $|\Gamma_a(i)| \geq 3$ is colored black;

- **Iterative coloring**: Start with $k = 1$

    1 Every agent with $\left|\Gamma_a(i)\right| + |\hat{\Gamma}_u^{(k-1)}(i)| \geq 3$, where $\hat{\Gamma}_u^{(k)}(i)$ is the set $\hat{\Gamma}_u(i)$ at step k, is colored black;
    2 Every agent j updates is own $\hat{\Gamma}_u^{(k)}(j)$ with the new colored nodes;
    3 The set $B_u^{(k)}$ is updated, where $B_u^{(k)}$ is the set B_u at step k;
    4 Set $k \leftarrow k + 1$ and repeat the previous steps until V_u contains only black nodes.

Suppose that the previous algorithm can color the entire graph black in a finite amount of steps, say n. Then, n is called RDML-lifetime.

Definition 2 (RDML-lifetime) *A directed graph G is said to have RDML-lifetime equal to n if and only if the RDML-coloring algorithm colors black the set V_u in exactly n steps. If no such integer exists, by convention, $n = +\infty$.*

This allows us to formally define compatibility:

Definition 3 (RDML-compatibility) *A directed graph G is said to be RDML-compatible if and only if:*

    1 *G is RDML-initializable;*
    2 *the RDML-lifetime of G is finite.*

Otherwise, G is said to be RDML-incompatible.

In practice, there are only two ways for which a graph is RDML-incompatible: either G cannot be initialized, or the RDML-lifetime of G is infinite. Testing the first condition is trivial; the interesting result is that testing the second condition is made simple thanks to the following.

Theorem 1 *An RDML-initializable graph G is RDML-incompatible if and only if there exist an integer h such that*

$$B_u^{(h)} = B_u^{(h-1)} \wedge \left|B_u^{(h)}\right| < |V_u| \tag{29}$$

that is, if there is a step h in which no more agents can be colored black and at least one agent is still left white.

Proof $(\Leftarrow)$ First, observe that, by construction, $B_u^{(k-1)} \subseteq B_u^{(k)}$, as black nodes can only be added. Let $C_G : \mathbb{N} \to \mathbb{N}$ be the following function

$$C_G(k) = \left|B_u^{(k)}\right|. \tag{30}$$

Since $B_u^{(k-1)} \subseteq B_u^{(k)}$, C_G is non-decreasing. An RDML-initializable graph is RDML-compatible if and only if it has finite RDML-lifetime, i.e., there must exist n such that $C_G(n) = |V_u|$. But condition (29) implies that there exists h such that $C_G(h) = C_G(h - 1) < |V_u|$. At step $h + 1$ and all successive steps, C_G cannot increase since the set $B_u^{(k)}$ cannot change. To show this, the key observation is that, as the graph G at step $h - 1$ did not satisfy the conditions for $B_u^{(h-1)}$ to grow (by hypothesis), the set was equal to itself at step h, i.e., $B_u^{(h-1)} = B_u^{(h)}$. But since no color change happened in V_u at step h, the graph G still does not satisfy the conditions for $B_u^{(k)}$ to grow for $k \geq h$. Thus, $C_G(k)$ becomes a constant function for $k \geq h$ and can never reach the value $|V_u|$.

$(\Rightarrow)$ Since G is an RDML-incompatible graph by hypothesis, at least one agent must be white, so $\left|B_u^{(h)}\right| < |V_u|$ is true for any h. Since $C_G(k)$ is non-decreasing, it must become constant for some $h > k$, but this implies that, for some h, $\left|B_u^{(h)}\right| = \left|B_u^{(h-1)}\right|$. This implies that,

since $B_u^{(k-1)} \subseteq B_u^{(k)}$ by construction, $B_u^{(h)} = B_u^{(h-1)}$ for some h. □

Definition 4 (RDML-depth) *Let G be a directed graph. Then,*

$$h_G \triangleq \inf_h \left\{ h \in \mathbb{N} : B_u^{(h)} = B_u^{(h+1)} \right\} \qquad (31)$$

is called the RDML-depth of G.

A complete graph has $h_G = 0$, as all agents are colored black during the initialization phase of the RDML-coloring algorithm.

Corollary 1 *Let G be a directed graph. Then, $h_G \leq n$, where n is the RDML-lifetime of G.*

Proof If G is not RDML-initializable, $h_G = 0$ as $B_u^{(0)} = \emptyset$. If G is RDML-initializable, there are two cases: either n is finite or not. In the latter, $n = +\infty$ and h_G is finite by previous theorem. If n is finite, $h_G = n$ since $\left| B_u^{(n)} \right| = |V_u|$ by definition of RDML-lifetime. □

The previous corollary proves that h_G is always finite, regardless of G. This allows us to write the graph compatibility test, shown in Algorithm 3. Thanks to the previous results, this algorithm always converges and can be used to test a priori if a graph is RDML-compatible or not.

Remark Algorithm 3 can be intuitively explained via a physical metaphor, where, in a metal grid (representing the graph), "heat" (information) spreads out starting from some initial "hot spots" (nodes that are colored black in the first iteration). This spreading is continued, reaching more and more locations on the grid, until the event occurs that further spreading of "heat" does not change the "heat map." If, at this point, there are cold spots (nodes that have not been colored black), the graph is RDML-incompatible. By contrast, if heat spreads throughout the grid, the graph is RDML-initializable.

Figures 1 and 2 show two examples of RDML-initializable graphs. Figure 1 illustrates the case of a small ($N_a = 4, N_u = 4$) but highly connected network. In contrast, Fig. 2 displays a larger ($N_a = 15, N_u = 20$), but weakly connected network (graph depth = 7). In particular, only 90 out of 680 possible directed links are connected, where anchors have three to four links (out of 20). For both cases, graph compatibility can be shown using Algorithm 3.

4 Results and discussion

In this section, we use standard Monte Carlo (MC) simulation to evaluate the performance of the proposed algorithm and its competitors. As a performance metric,

---

Algorithm 3 Graph Compatibility Test

Input: Directed Graph G with $V = V_a \cup V_u$
Output: Binary flag f_G, Depth h_G
    Initialization :
1: Set $Color(V_u) \leftarrow 0$; %*white*
2: **for** $i = 1$ to $|V_u|$ **do**
3:    **if** $|\Gamma_a(i)| \geq 3$ **then**
4:      $Color(i) \leftarrow 1$; %*black*
5:    **end if**
6: **end for**
7: Initialize set $\hat{\Gamma}_u^{(0)}(i)$ for every agent i and set $B_u^{(0)}$;
8: Compute $C_G(0) = \left| B_u^{(0)} \right|$; **if** $C_G(0) = |V_u|$, **stop** and **return** $(f_G, h_G) = (1, 0)$;
    Coloring Loop
9: Set $k \leftarrow 1$;
10: **while** (1) **do**
11:    **for** $i = 1$ to $|V_u|$ **do**
12:      **if** $|\Gamma_a(i)| + |\hat{\Gamma}_u^{(k-1)}(i)| \geq 3$ **then**
13:        $Color(i) \leftarrow 1$;
14:      **end if**
15:    **end for**
16:    Update $\hat{\Gamma}_u^{(k)}(i)$ for every agent i and compute $C_G(k) \leftarrow |B_u^{(k)}|$;
    Exit condition
17:    **if** $C_G(k-1) = C_G(k)$ **then**
18:      $h_G \leftarrow k - 1$;
19:      **if** $C_G(k) < |V_u|$ **then**
20:        $f_G \leftarrow 0$; %*graph G is RDML-incompatible*
21:      **else**
22:        $f_G \leftarrow 1$; %*graph G is RDML-compatible*
23:      **end if**
24:      **break** loop;
25:    **end if**
26:    $k \leftarrow k + 1$;
27: **end while**
28: **return** f_G, h_G.

---

we will show the ECDF (empirical cumulative distribution function) of the localization error, defined as $e_i \triangleq \|\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i\|$ for agent i, i.e., an estimate of the probability $\mathcal{P}\left\{ \|\hat{\boldsymbol{x}}_i - \boldsymbol{x}_i\| \leq \gamma \right\}$. The ECDFs are obtained by stacking all the localization errors for every agent in a single vector, in order to give a global picture of the algorithm performances. The simulated scenario is as follows. In a square coverage area of 100×100 m^2, $N_a = 11$ stationary anchors are deployed, as depicted in Fig. 3. The channel parameters are generated as follows: $p_{0_{\text{LOS}}} \sim \mathcal{U}[-30, 0]$, $\alpha_{\text{LOS}} \sim \mathcal{U}[2, 4]$, and $\sigma_{\text{LOS}} = 6$, while, for the NLoS case, $p_{0_{\text{NLOS}}} \sim \mathcal{N}(0, 25)$, $\alpha_{\text{NLOS}} \sim \mathcal{U}[3, 6]$, and $\sigma_{\text{NLOS}} = 12$. Similar settings on the reference power and the path loss exponent can be found in [46, 47], respectively. At each
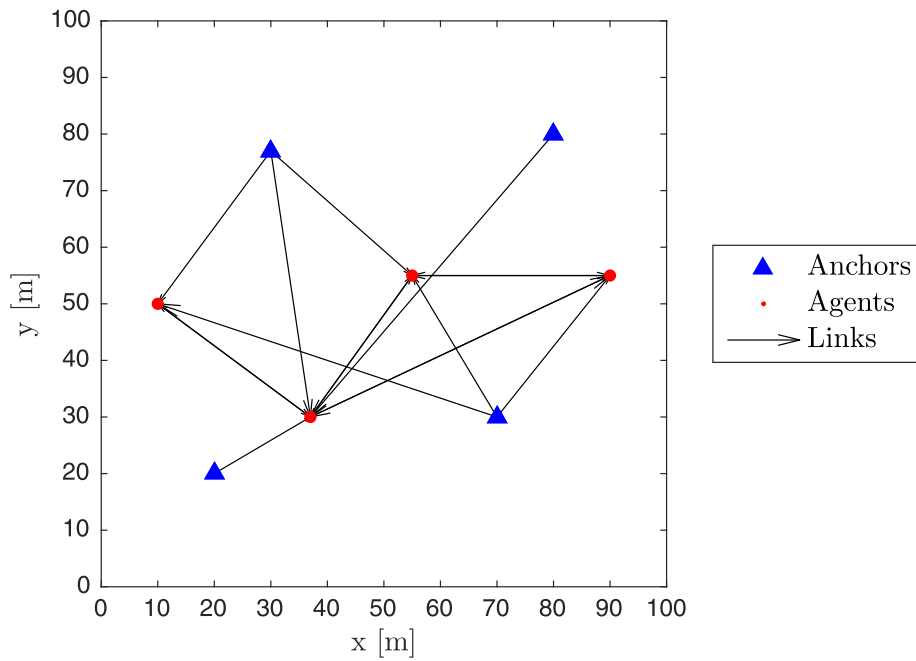
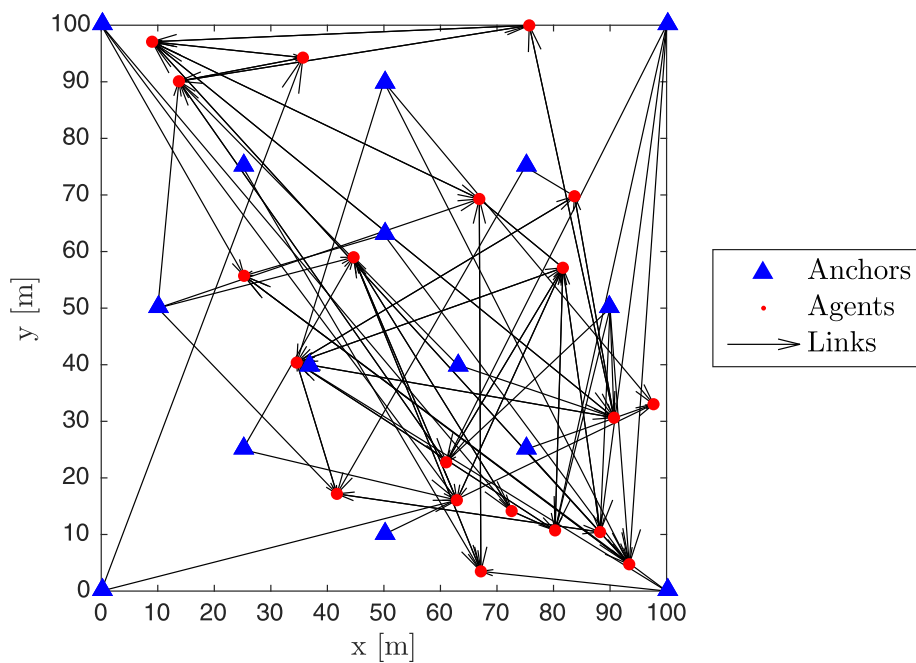**Fig. 1** Example of a compatible graph with $N_a = 4, N_u = 4$

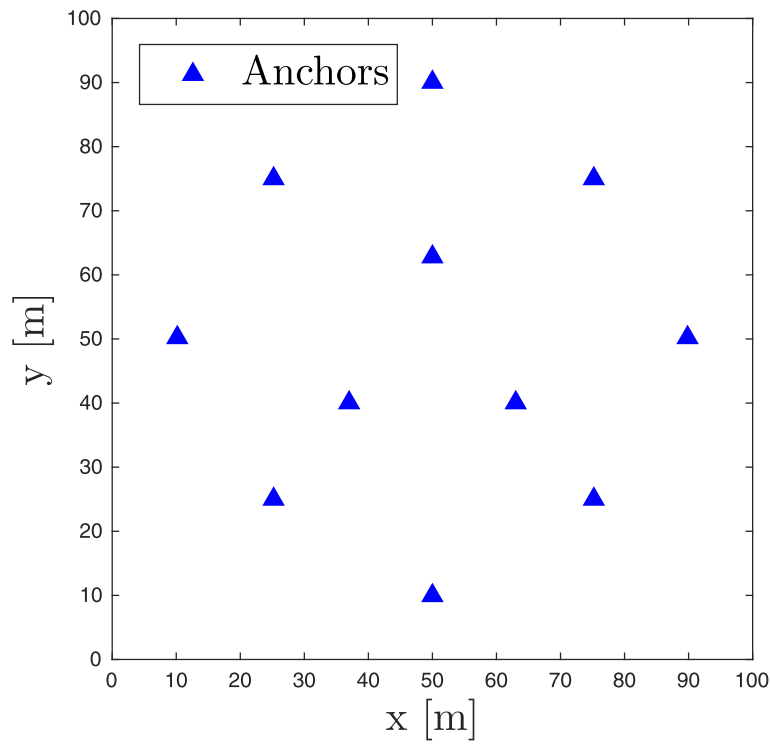**Fig. 2** Example of a compatible graph with $N_a = 15, N_u = 20$

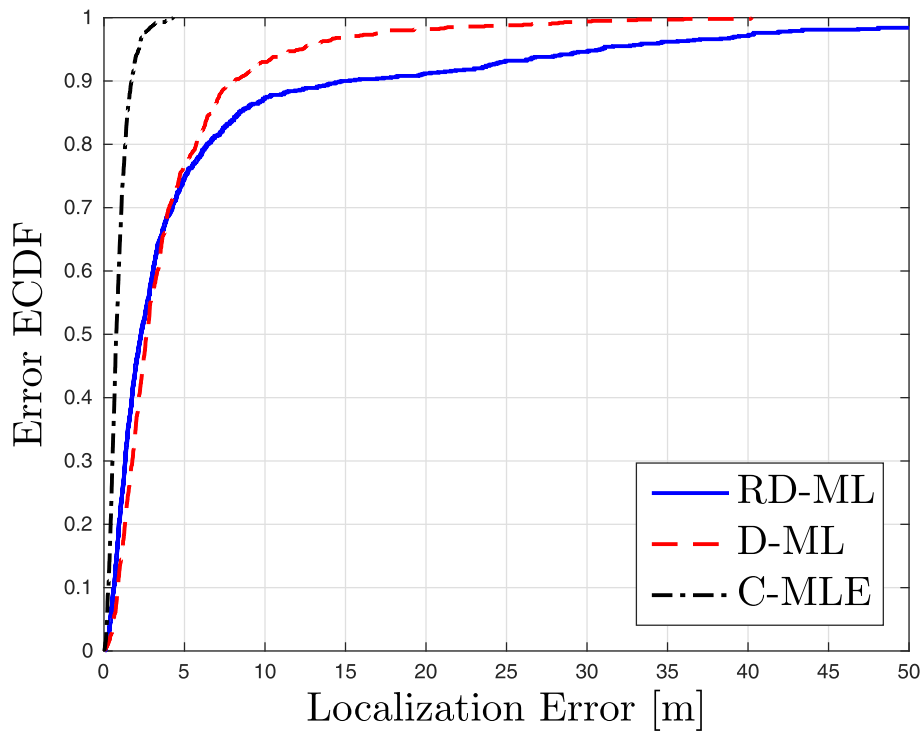**Fig. 3** Anchors positions used in all simulations

Fig. 4 All LoS links with white Gaussian noise. Fully connected networks, $N_a = 11, N_u = 10$

MC trial, $N_u = 10$ agents are randomly (uniformly) generated in the coverage area. Unless stated otherwise, $K = 40$ samples per link are used. Finally, each simulation consists of 100 MC trials.

The optimization problems (27) and (19) have been solved as follows. For DML, a 2D `grid search` has been used, while, for RDML and C-MLE, the optimization has been performed with the MATLAB solver `fmincon`.

4.1 Gaussian noise

Here, we validate our novel approach by considering fully connected networks in three different scenarios. In Fig. 4, all links are assumed to be LoS. As can be seen from the ECDFs, the robust approach has similar performances to the D-ML algorithm, which was designed assuming all links to be of the same type. In Fig. 5, all links are assumed to be NLoS and again we see a similar result. The key difference is in Fig. 6, where a mixed LoS/NLoS scenario is considered and the fraction of NLoS links is random (numerically generated as $\mathcal{U}[0, 1]$ and different for each MC trial). Here, the robust approach outperforms D-ML, validating our strategy. Moreover, it has a remarkably close performance to the centralized algorithm C-MLE. As a result, the robust approach is the preferred option in all scenarios.

4.2 Non-Gaussian noise for NLoS

In order to evaluate robustness, we consider a model mismatch on the noise distribution by choosing, for the NLoS links, a Student*t* distribution with $\nu = 5$ degrees of freedom (as usual, serially i.i.d. and independent from the LoS noise sequence). For brevity, we consider only the case of a fully connected network in a mixed LoS/NLoS scenario. As shown by Fig. 7, our proposed approach is able to cope with non-Gaussianity without a significant performance loss, thanks to the Gaussian mixture model.

4.3 Cooperative gain

The proposed algorithm exhibits the so called "cooperative gain," i.e., a performance advantage (according to some performance metric, typically localization accuracy) with respect to a non-cooperative approach, where each node tries to localize itself only by self-localization (12). In our case, the cooperative gain is twofold: first, it allows to improve localization accuracy; second, it allows to localize otherwise non-localizable agents. To show the first point, we consider a mixed LoS/NLoS environment (for simplicity, with white Gaussian noise) and a network is generated at each Monte Carlo trial, assuming a communication radius $R = 70$ m, with an ideal probability of detection model [17]. Moreover, the network is generated in a way as to allow all agents to self-localize and, as a
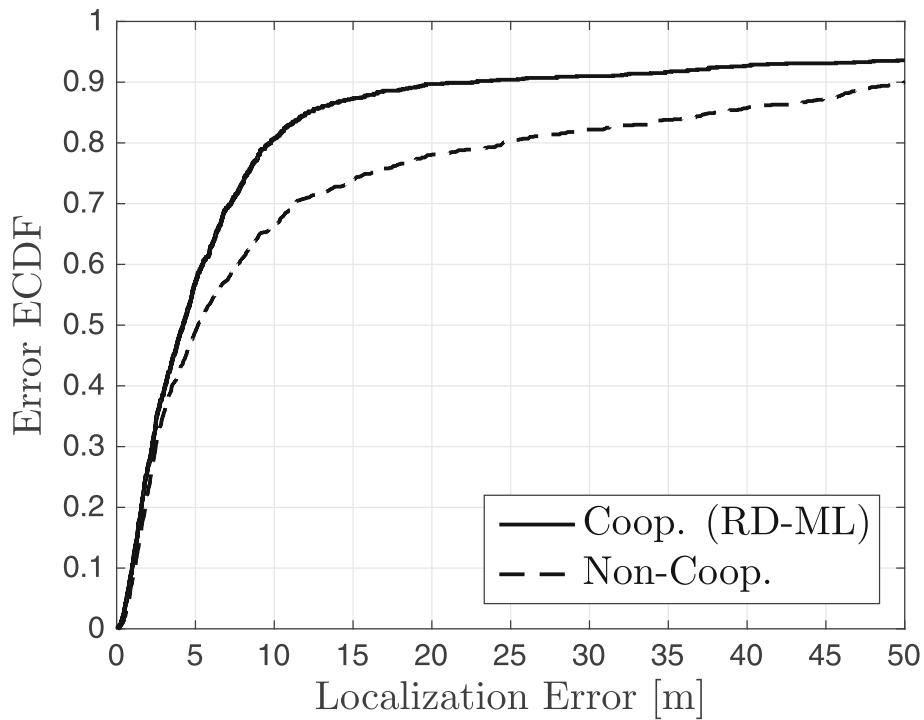

Fig. 5 All NLoS links with white Gaussian noise. Fully connected networks, $N_a = 11, N_u = 10$

Fig. 6 Mixed LoS/NLoS scenario with white Gaussian noise. Fully connected networks, randomized NLoS fraction, $N_a = 11, N_u = 10$

Fig. 7 Mixed LoS/NLoS scenario with Student's t distributed noise with $\nu = 5$ under NLoS. Fully connected networks, randomized NLoS fraction, $N_a = 11, N_u = 10$
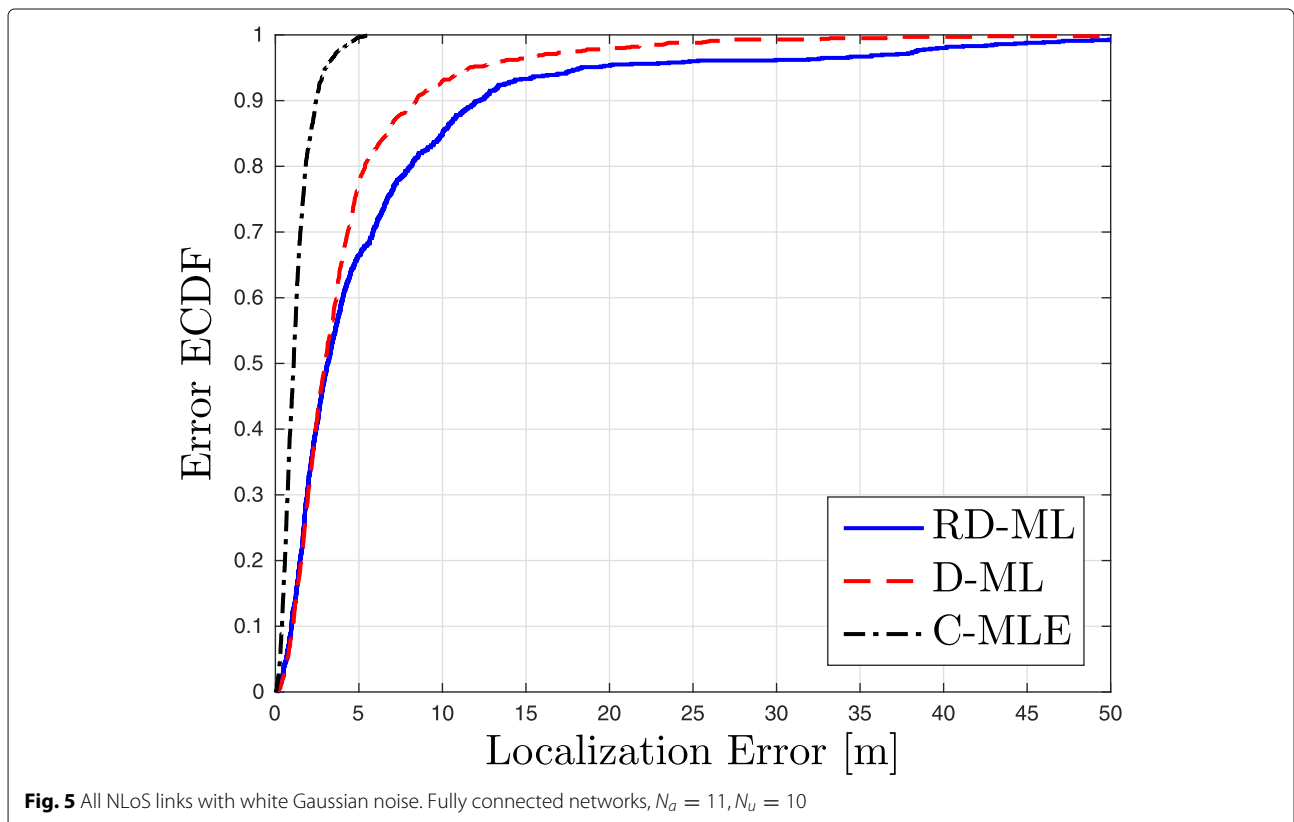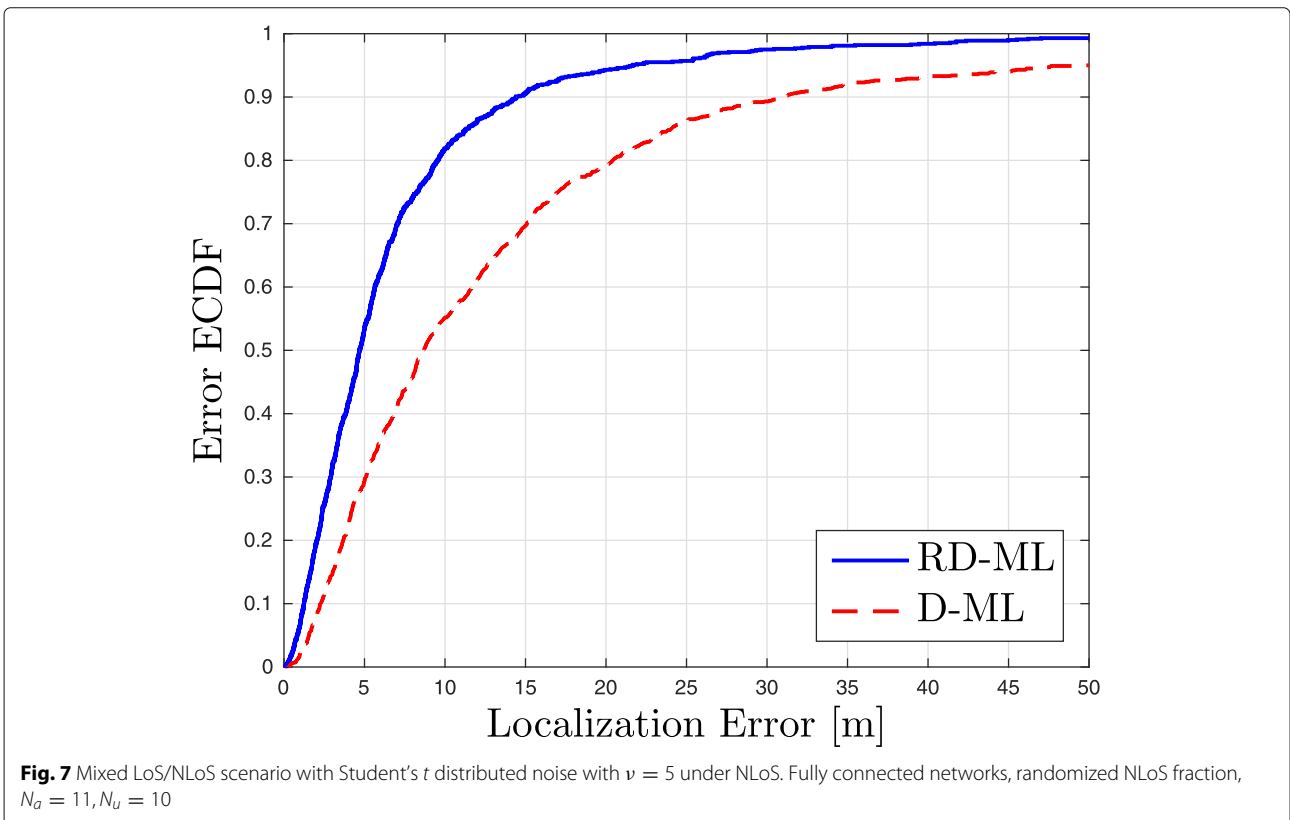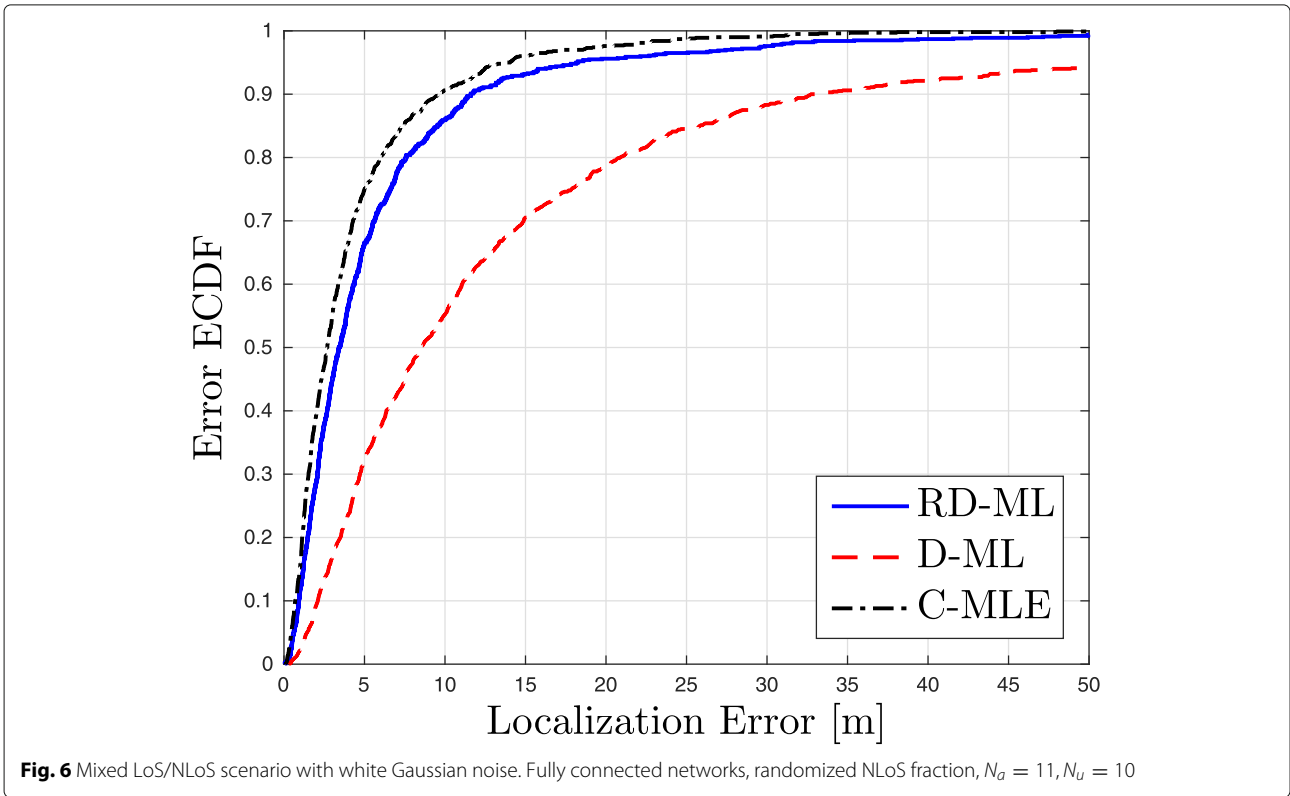
Fig. 8 Cooperative (RDML) and non-cooperative approaches in a mixed LoS/NLoS scenario, in a network with communication radius $R = 70$ m, $N_a = 11, N_u = 10$
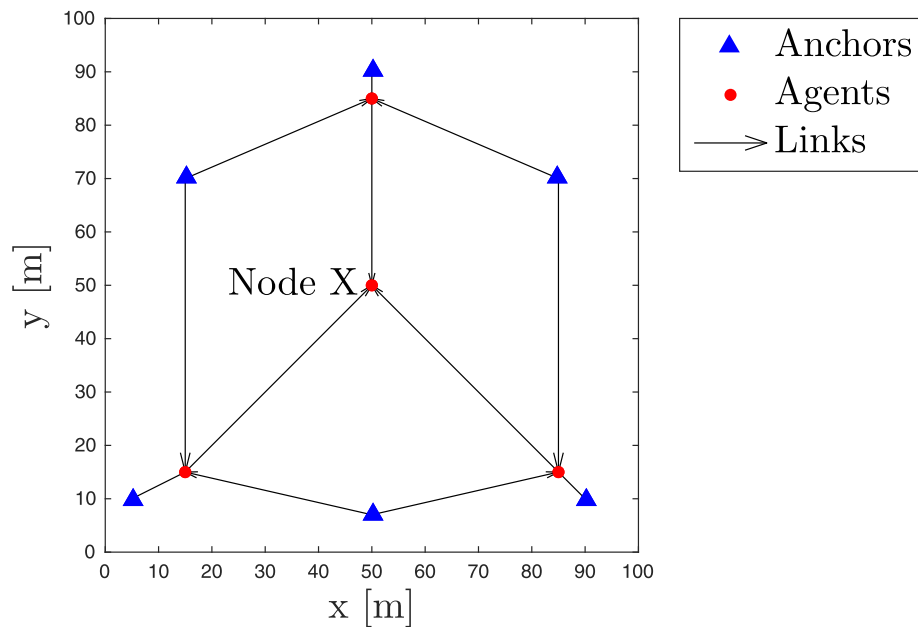

Fig. 9 Network of a toy example, with $N_a = 6, N_u = 4$

consequence, it is RDML-compatible. In Fig. 8, the ECDFs of the localization error are shown. To show the second point, we consider the toy network depicted in Fig. 9. In this example, agent X is not able to self-localize, since no anchors are in its neighborhood, $\Gamma_a(X) = \emptyset$. Thus, in a non-cooperative approach, the position of agent X cannot be uniquely determined[7], while, in a cooperative approach, the position of agent X can actually be obtained by exchanging information, e.g., the estimated positions of its neighbors.

4.4 Variable *K* and NLoS fraction

We next analyze the RDML algorithm by varying the number of samples per link, K, and the fraction of NLoS links. In both cases, we consider (for brevity) fully connected networks with white Gaussian noise; for variable K, the fraction of NLoS is randomized, while, for variable NLoS fraction, K is fixed. In Fig. 10, we observe that the performance increases as K increases, as expected. In Fig. 11, where each point represents 100 MC trials, the median error is chosen as a performance metric[8] and RDML shows good performances and is not significantly affected by the actual NLoS fraction, which is evidence for its robustness, while D-ML is clearly inferior and suffers from model mismatch.

4.5 Communication and computational costs

The communication overhead is evaluated by computing the number of *elementary messages* sent throughout the network, where an elementary message is simply defined as a single scalar. Sending a d-dimensional message, e.g., (x, y) coordinates, is equivalent to sending d elementary messages. In Fig. 12, the communication overhead of RDML and DML is plotted with respect to the number of agents, assuming a complete graph: for N_u agents, $2N_u(N_u - 1)$ scalars, representing estimated 2D positions, are needed for RDML, while $4N_u(N_u - 1)$ are needed for DML, as estimates of (p_0, α) are also necessary. Viewing (2D) position has the fundamental information of a message ($d = 2$), in a complete graph, RDML achieves the theoretical minimum cost in order for all nodes to have complete information, while DML uses auxiliary information. For a general graph, the communication overhead depends on the specific graph topology, but is never greater than the aforementioned value. Thus, for a general graph, the communication cost is upper-bounded by a quadratic function of N_u.

Regarding computational complexity, both RDML and DML scale linearly with N_u and they benefit from parallelization, as the main optimization task can be executed independently for each involved node. As already
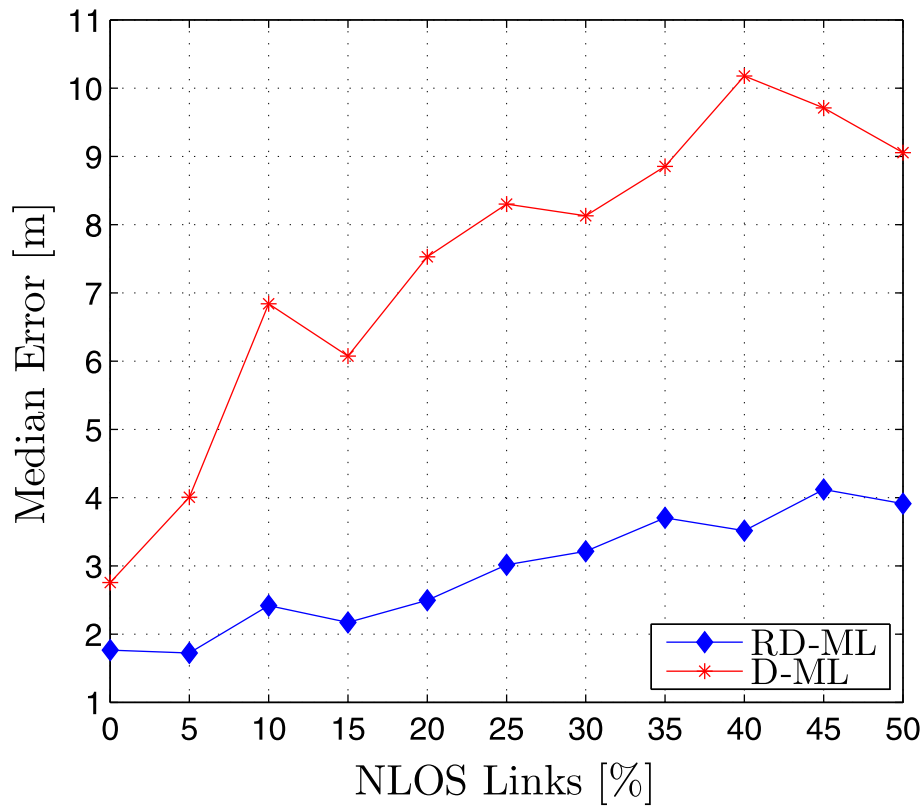

Fig. 10 Error ECDFs of RDML for various values of *K* in fully connected networks with white Gaussian noise and randomized NLoS fraction, $N_a = 11, N_u = 10$
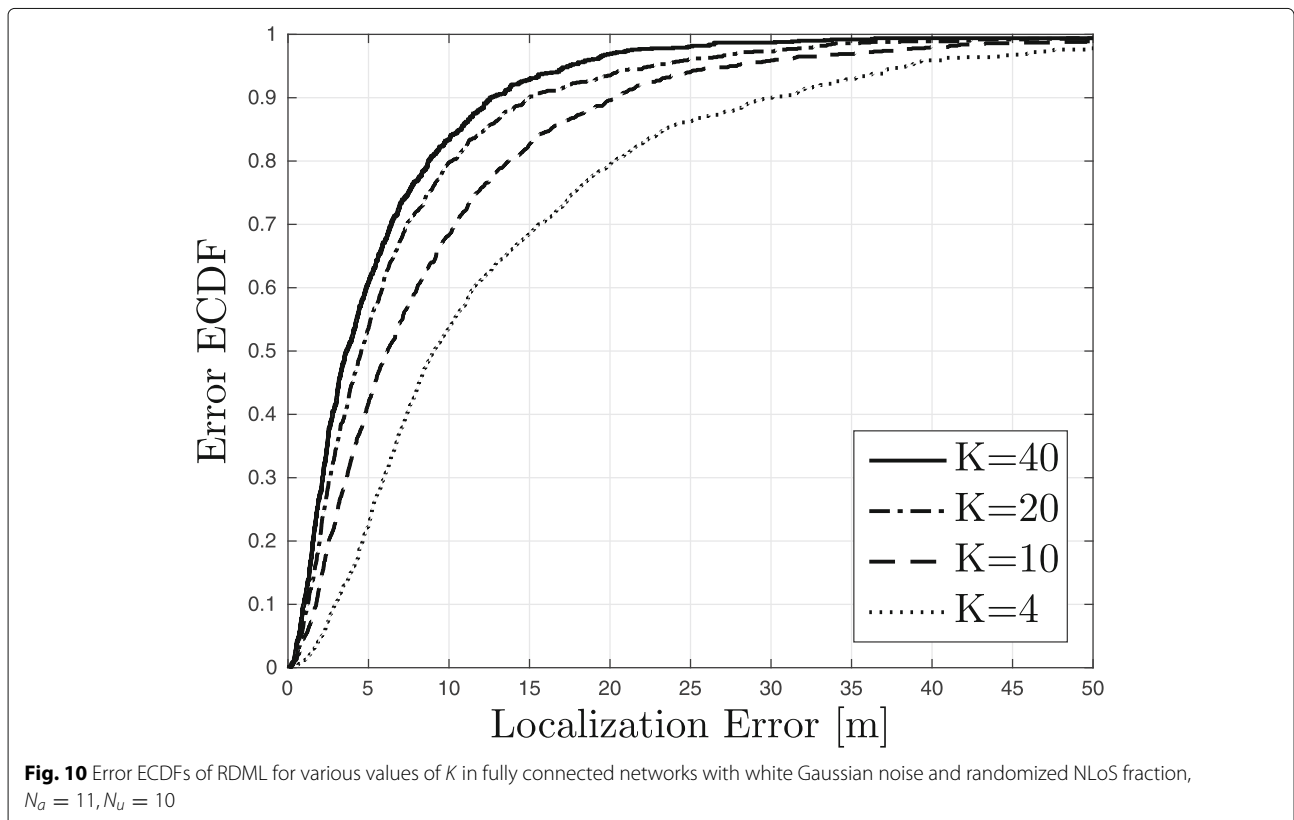
Fig. 11 Median error of RDML and D-ML for a variable fraction of NLoS links. Fully connected networks, $K = 40$, $N_a = 11$, $N_u = 10$
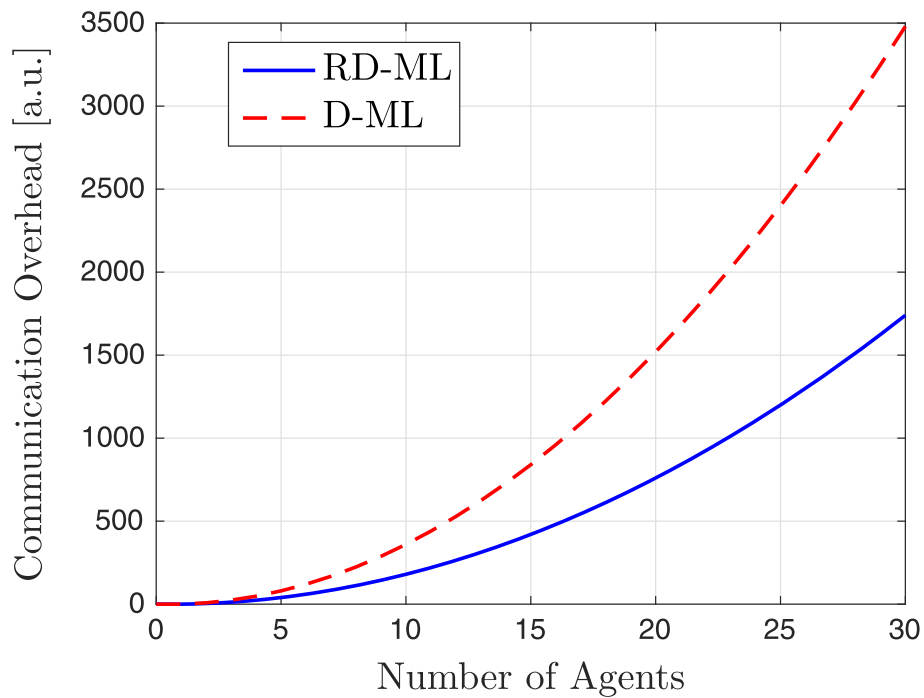

Fig. 12 Communication costs for a fully-connected network

mentioned, DML operates a trade-off between communication and computational complexity; in fact, the DML optimization problem (27) is easier to solve than the RDML optimization problem (19). Both problems are non-convex and may have local minima/maxima, so care must be taken in the optimization procedure.

5 Conclusions

We have developed a novel, robust ML-based scheme for RSS-based distributed cooperative localization in mixed LoS/NLoS scenarios, which, while not being optimal, has good overall accuracy, is adaptive to the environment changes, is robust to NLoS propagation, including non-Gaussian noise, and has communication overhead upper-bounded by a quadratic function of the number of agents and computational complexity scaling linearly with the number of agents, also benefiting from parallelization. The main original contributions are that, unlike many algorithms present in the literature, the proposed approach (a) does not require calibration and (b) does not require symmetrical links (nor does it resort to symmetrization), thus naturally accounting for the topology of directed graphs with asymmetrical links as a result of miss-detections and channel anisotropies. To the best of the authors' knowledge, this is the first distributed cooperative RSS-based algorithm for *directed* graphs. The main disadvantage is imposing some restrictions on the arbitrariness of networks' topology, but these restrictions disappear for sufficiently connected networks. We also derive a compatibility test based on graph coloring, which allows to determine whether the given network is compatible. If it is compatible, convergence of the algorithm is guaranteed. Future work may include the consideration of possible approximations, in order to extend this approach to more general networks, and of alternative models, overcoming the limitations of the standard path loss model.

Endnotes

[1] Throughout the paper, vectors and matrices will be denoted in bold, $\|\boldsymbol{v}\|$ denotes the Euclidean norm of vector $\boldsymbol{v}$, $|\mathcal{A}|$ denotes the cardinality of set $\mathcal{A}$. We denote by $\mathbb{E}[X]$ and $\mathrm{Var}[X]$ the statistical expectation and variance, respectively, of random variable X. Finally, $\mathbb{B} = \{0, 1\}$ is the Boolean set.

[2] The values on the main diagonal are arbitrary. Here we choose $m_{i \to i} = 1$.

[3] $\delta_{i,j} = 1$ if and only if $i = j$, zero otherwise.

[4] For better readability, the notation $^{(m)}$ has not been carried over, as it implicit in the formalism.

[5] The reason for this is that localizing a node in 2D requires at least three anchors.

[6] Hereafter, we omit the conditioning on the set $\{o_{n \to i}\}$ of actually observed RSS measures (received by node i) in the joint likelihood function, since it is implicit in the neighborhood formalism.

[7] This easily follows by observing that the deterministic (noiseless) version of all the relevant equations for agent X admit infinite solutions.

[8] This is due to the fact that RSME (Root Mean Square Error) is not a suitable metric when the Error ECDFs are very long-tailed, as in our case.

6 Appendix A: On using the time-averaged sample mean

Let $\Theta \subseteq \mathbb{R}^d$ be the (non-empty) parameter space, $\boldsymbol{\theta} \in \Theta$ unknown deterministic parameters and $s : \Theta \to \mathbb{R}$ a function. Consider the following model

$$y(k) = s(\boldsymbol{\theta}) + w(k), \quad w(k) \sim \mathcal{N}\left(0, \sigma^2\right) \tag{32}$$

where $k = 1, \ldots, K$ and the noise sequence $w(k)$ is zero-mean, i.i.d. Gaussian, with deterministic unknown variance $\sigma^2 > 0$. Let

$$\bar{y} = \frac{1}{K} \sum_{k=1}^{K} y(k) \tag{33}$$

be the sample mean and let $p\left(y(1), \ldots, y(k); \boldsymbol{\theta}, \sigma^2\right)$ denote the joint likelihood function. Then,

$$\arg\max_{\boldsymbol{\theta} \in \Theta} p\left(y(1), \ldots, y(k); \boldsymbol{\theta}, \sigma^2\right) = \arg\max_{\boldsymbol{\theta} \in \Theta} p\left(\bar{y}; \boldsymbol{\theta}, \sigma^2\right) \tag{34}$$

Proof Since the observations are independent,

$$p\left(y(1), \ldots, y(k); \boldsymbol{\theta}, \sigma^2\right) = \prod_{k=1}^{K} p\left(y(k); \boldsymbol{\theta}, \sigma^2\right)$$

$$= \frac{1}{(2\pi\sigma^2)^{K/2}} e^{-\frac{1}{2\sigma^2} \sum_{k=1}^{K} (y(k) - s(\boldsymbol{\theta}))^2} \tag{35}$$

We can now focus on the term in the exponential. By adding and subtracting $\bar{y}$ from the quadratic term and expanding it, we get

$$Q = \sum_{k=1}^{K} (y(k) - s(\boldsymbol{\theta}))^2 = \sum_{k=1}^{K} ((y(k) - \bar{y}) - (s(\boldsymbol{\theta}) - \bar{y}))^2$$

$$= \sum_{k=1}^{K} (y(k) - \bar{y})^2 + \sum_{k=1}^{K} (s(\boldsymbol{\theta}) - \bar{y})^2 - 2 \sum_{k=1}^{K} (y(k) - \bar{y})(s(\boldsymbol{\theta}) - \bar{y}) \tag{36}$$

Observing now that

$$\sum_{k=1}^{K} (y(k) - \bar{y})\,(s(\boldsymbol{\theta}) - \bar{y}) = \sum_{k=1}^{K} \left(y(k)s(\boldsymbol{\theta}) - y(k)\bar{y} - \bar{y}s(\boldsymbol{\theta}) + \bar{y}^2 \right)$$
$$= 0 \tag{37}$$

we are left with

$$Q = \sum_{k=1}^{K} (y(k) - \bar{y})^2 + \sum_{k=1}^{K} (s(\boldsymbol{\theta}) - \bar{y})^2 \tag{38}$$

Thus, the joint pdf can be factorized as follows

$$p\left(y(1), \ldots, y(k); \boldsymbol{\theta}, \sigma^2\right)$$
$$= \underbrace{\frac{1}{\left(2\pi\sigma^2\right)^{K/2}} e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{K}(y(k)-\bar{y})^2}}_{=h(y(1),\ldots,y(k);\sigma^2)} \underbrace{e^{-\frac{1}{2\sigma^2}\sum_{k=1}^{K}(s(\boldsymbol{\theta})-\bar{y})^2}}_{=g(\bar{y};\boldsymbol{\theta},\sigma^2)} \tag{39}$$

As a by-product, by invoking the Neyman-Fisher factorization theorem [45] and assuming known σ^2, $\bar{y}$ is a sufficient statistic for $\boldsymbol{\theta}$. Resuming our proof, we can now observe that

$$\arg\max_{\boldsymbol{\theta}\in\Theta} p\left(y(1), \ldots, y(k); \boldsymbol{\theta}, \sigma^2\right) = \arg\max_{\boldsymbol{\theta}\in\Theta} g\left(\bar{y}; \boldsymbol{\theta}, \sigma^2\right)$$
$$= \arg\max_{\boldsymbol{\theta}\in\Theta} \left\{ -\sum_{k=1}^{K}(s(\boldsymbol{\theta}) - \bar{y})^2 \right\} = \arg\min_{\boldsymbol{\theta}\in\Theta}(\bar{y} - s(\boldsymbol{\theta}))^2 \tag{40}$$

since no term in the summation depends on k. Thus,

$$\arg\max_{\boldsymbol{\theta}\in\Theta} p\left(y(1), \ldots, y(k); \boldsymbol{\theta}, \sigma^2\right) = \arg\min_{\boldsymbol{\theta}\in\Theta}(\bar{y} - s(\boldsymbol{\theta}))^2 \tag{41}$$

On the other hand,

$$\arg\max_{\boldsymbol{\theta}\in\Theta} p\left(\bar{y}; \boldsymbol{\theta}, \sigma^2\right)$$
$$= \arg\max_{\boldsymbol{\theta}\in\Theta} \frac{1}{\sqrt{2\pi\frac{\sigma^2}{K}}} \exp\left\{ -\frac{K}{2\sigma^2}(\bar{y} - s(\boldsymbol{\theta}))^2 \right\}$$
$$= \arg\max_{\boldsymbol{\theta}\in\Theta} \left\{ -\frac{K}{2\sigma^2}(\bar{y} - s(\boldsymbol{\theta}))^2 \right\}$$
$$= \arg\min_{\boldsymbol{\theta}\in\Theta}(\bar{y} - s(\boldsymbol{\theta}))^2 \tag{42}$$

which completes the proof. □

7 Appendix B: C-MLE derivation

The MLE of $\boldsymbol{x}$ is given by

$$\hat{\boldsymbol{x}}_{ML} = \arg\max_{\boldsymbol{x}} p(\Upsilon; \boldsymbol{x}) \tag{43}$$

where $p(\Upsilon; \boldsymbol{x})$ is the joint likelihood function. Since all other parameters are assumed known, the set of all time-averaged measures $\bar{\Upsilon} = \cup_{i=1}^{N_u} \bar{\mathcal{Y}}_i$ is a sufficient statistic for $\boldsymbol{x}$ (see Appendix A), from which it follows that

$$\arg\max_{\boldsymbol{x}} p(\Upsilon; \boldsymbol{x}) = \arg\max_{\boldsymbol{x}} p\left(\bar{\Upsilon}; \boldsymbol{x}\right) \tag{44}$$

By link independence,

$$p(\bar{\Upsilon}; \boldsymbol{x}) = \prod_{i=1}^{N_u} \prod_{j\in\Gamma(i)} p(\bar{r}_{j\to i}; \boldsymbol{x}) \tag{45}$$

where $p(\bar{r}_{j\to i}; \boldsymbol{x})$ is the marginal likelihood function. Thus, we have

$$\hat{\boldsymbol{x}}_{ML} = \arg\max_{\boldsymbol{x}} \prod_{i=1}^{N_u} \prod_{j\in\Gamma(i)} \frac{1}{\left(2\pi\sigma_{j\to i}^2\right)}$$
$$\times\, e^{-\frac{1}{2\sigma_{j\to i}^2}(\bar{r}_{j\to i} - p_{0_{j\to i}} + 10\alpha_{j\to i}\log_{10}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|)^2} \tag{46}$$

Taking the natural logarithm and neglecting constants,

$$\hat{\boldsymbol{x}}_{ML}$$
$$= \arg\min_{\boldsymbol{x}} \sum_{i=1}^{N_u} \sum_{j\in\Gamma(i)} \frac{1}{\sigma_{j\to i}^2} (\bar{r}_{j\to i} - p_{0_{j\to i}} + 10\alpha_{j\to i}\log_{10}\|\boldsymbol{x}_j - \boldsymbol{x}_i\|)^2 \tag{47}$$

which completes the derivation.

Authors' information

LC received his B.Sc. degree with full marks in Information Engineering at University of Salento, Lecce, Italy, in 2012. He received his M.Sc. degree in Telecommunications Engineering (*summa cum laude*) at University of Salento, in 2014. He started his Ph.D. studies at University of Salento in 2014 and finished his Ph.D. in 2018. His main area of research is signal processing, while his research topic is applying signal processing techniques to the problem of localization, with a specific focus on localization based upon received signal strength. His other research interests include radar applications, target tracking, data mining and information theory.

DJ received the B.Sc. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2011, and the M.Sc. degree in electrical engineering and information technology from Technische Universität Darmstadt, Darmstadt, Germany, in 2014. She is currently working towards the Ph.D. degree in the Signal Processing Group at Technische Universität Darmstadt. Her research interests include localization and tracking, distributed and cooperative inference in wireless networks.

MM received the Dipl.-Ing (2009) and the Dr.-Ing. degree (*summa cum laude*) in Electrical Engineering and Information Technology (2014), both from Technische Universität Darmstadt, Darmstadt, Germany. He completed his diploma thesis with the Contact Lens and Visual Optics Laboratory, School of Optometry, Brisbane, Australia, on the role of cardiopulmonary signals in the dynamics of the eye's wavefront aberrations. Currently, he is a Postdoctoral Fellow at the Signal Processing Group, Institute of Telecommunications and has recently been awarded Athene Young Investigator of Technische Universität Darmstadt. His research is on robust statistics for signal processing with applications in biomedical signal processing, wireless sensor networks, and array signal processing. MM was the supervisor of the Technische Universität Darmstadt student team who won the international IEEE Signal Processing Cup 2015. MM co-organized the 2016 Joint IEEE SPS and EURASIP Summer School on Robust Signal Processing. In 2017, together with his coauthors, MM received the IEEE Signal Processing Magazine Best Paper Award for the paper entitled "Robust Estimation in Signal Processing: A tutorial-style treatment of fundamental concepts". In 2017, he was elected to the European Association for Signal Processing (EURASIP) Special Area Team in Theoretical and Methodological Trends in Signal Processing (SAT-TMSP).

AZ is a Fellow of the IEEE and IEEE Distinguished Lecturer (Class 2010- 2011). He received his Dr.-Ing. from Ruhr-Universität Bochum, Germany in 1992. He was with Queensland University of Technology, Australia from 1992–1998 where he was Associate Professor. In 1999, he joined Curtin University of Technology, Australia as a Professor of Telecommunications. In 2003, he moved to Technische Universität Darmstadt, Germany as Professor of Signal Processing and Head of the Signal Processing Group. His research interest lies in statistical methods for signal processing with emphasis on bootstrap techniques, robust detection and estimation and array processing applied to telecommunications, radar, sonar, automotive monitoring and safety, and biomedicine. He published over 400 journal and conference papers on the above areas. AZ served as General Chair and Technical Chair of numerous international conferences and workshops; more recently he was the Technical Co-Chair of ICASSP-14 held in Florence, Italy. He also served on publication boards of various journals, notably as Editor-In-Chief of the IEEE Signal Processing Magazine (2012–2014). AZ was the Chair (2010–2011) of the IEEE Signal Processing Society (SPS) Technical Committee Signal Processing Theory and Methods (SPTM). He served on the Board of Governors of the IEEE SPS (2015–2017) and is the president of the European Association of Signal Processing (EURASIP) (2017–2018).

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Author details

[1]Department of Innovation Engineering, University of Salento, Lecce, Italy. [2]Signal Processing Group, Department of Electrical Engineering and Information Technology, Technische Universität Darmstadt, Darmstadt, Germany.

References

1. F. Gustafsson, F. Gunnarsson, Mobile positioning using wireless networks. IEEE Signal Process. Mag. **22**(4), 41–53 (2005)
2. N. Patwari, J. Ash, S. Kyperountas, A. O. I. Hero, R. Moses, N. Correal, Locating the nodes: Cooperative localization in wireless sensor networks. IEEE Signal Process. Mag. **22**(4), 54–69 (2005)
3. S. Gezici, A survey on wireless position estimation. Wireless Pers. Commun. **44**(3), 263–282 (2008)
4. L. M. Trevisan, M. E. Pellenz, M. C. Penna, R. D. Souza, M. S. Fonseca, A simple iterative positioning algorithm for client node localization in WLANs. EURASIP Wireless, J. Commun. and Netw. **2013**(1), 276 (2013)
5. A. Coluccia, F. Ricciato, A Software-defined radio tool for experimenting with RSS measurements in IEEE 802.15.4: implementation and applications. Int. Sensor J. Netw. **14**(3), 144–154 (2013)
6. T. S. Rappaport, *Wireless Communications: Principles and Practice (2nd Ed.)* (Prentice Hall, Upper Saddle River, NJ, USA, 2002)
7. J. Koo, H. Cha, Localizing WiFi access points using signal strength. IEEE Commun. Lett. **15**(2), 187–189 (2011)
8. A. Matic, A. Papliatseyeu, V. Osmani, O. Mayora-Ibarra, in *Proc. 8th IEEE Int. Conf. Pervasive Comput. Commun. (PerCom)*. Tuning to your position: FM radio based indoor localization with spontaneous recalibration, (Mannheim, Germany, 2010), pp. 153–161
9. M. Hellebrandt, R. Mathar, M. Scheibenbogen, Estimating position and velocity of mobiles in a cellular radio network. IEEE Trans. Veh. Technol. **46**(1), 65–71 (1997)
10. F. Yin, A. Li, A. M. Zoubir, C. Fritsche, F. Gustafsson, in *In Proc. 5th IEEE Int. Workshop Computat. Adv. Multi-Sensor Adapt. Process. (CAMSAP)*. RSS-based sensor network localization in contaminated Gaussian measurement noise, (2013), pp. 121–124
11. F. Bandiera, A. Coluccia, G. Ricci, A. Toma, in *Proc. 39th IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*. RSS-based localization in non-homogeneous environments, (Florence, Italy, 2014), pp. 4247–4251
12. F. Bandiera, A. Coluccia, G. Ricci, A Cognitive Algorithm for Received Signal Strength Based Localization. IEEE Trans. Signal Process. **63**(7), 1726–1736 (2015)
13. F. Bandiera, L. Carlino, A. Coluccia, G. Ricci, in *Proc. IEEE 6th International Workshop Comput. Adv. Multichannel Sensor Array Process. (CAMSAP)*. RSS-based localization of a moving node in homogeneous environments, (Cancun, Mexico, 2015), pp. 249–252
14. F. Bandiera, L. Carlino, A. Coluccia, G. Ricci, in *Proc. 19th Int. Conf. Inform. Fusion (FUSION)*. A cognitive algorithm for RSS-based localization of possibly moving nodes, (2016), pp. 2186–2191
15. S. Tomic, M. Beko, R. Dinis, Distributed RSS-AOA Based Localization With Unknown Transmit Powers. IEEE Wireless Commun. Lett. **5**(4), 392–395 (2016)
16. P. Abouzar, D. G. Michelson, M. Hamdi, RSSI-Based Distributed Self-Localization for Wireless Sensor Networks Used in Precision Agriculture. IEEE Trans. Wireless Commun. **15**(10), 6638–6650 (2016)
17. R. Zekavat, R. M. Buehrer, *Handbook of Position Location: Theory, Practice, And Advances.* (Wiley-IEEE press, Hoboken, 2011)
18. L. Y., F. S., G. D., C. K., in *Proc. 6th IEEE Int. Sympos. Microw., Ant., Prop., EMC Technol. (MAPE)*. A non-cooperative localization scheme with unknown transmission power and path-loss exponent, (2015), pp. 720–725
19. D. Jin, F. Yin, C. Fritsche, A. M. Zoubir, F. Gustafsson, in *Proc. 41st IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*. Cooperative localization based on severely quantized RSS measurements in wireless sensor network, (2016), pp. 4214–4218
20. S. Tomic, M. Beko, R. Dinis, RSS-Based Localization in Wireless Sensor Networks Using Convex Relaxation: Noncooperative and Cooperative Schemes. IEEE Trans. Veh. Technol. **64**(5), 2037–2050 (2015)
21. F. Yin, C. Fritsche, F. Gustafsson, A. M. Zoubir, in *Proc. 38th IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*. Received signal strength-based joint parameter estimation algorithm for robust geolocation in LOS/NLOS environments, (2013), pp. 6471–6475
22. F. Yin, C. Fritsche, D. Jin, F. Gustafsson, A. M. Zoubir, Cooperative localization in WSNs using Gaussian mixture modeling: Distributed ECM algorithms. IEEE Trans. Signal Process. **63**(6), 1448–1463 (2015)
23. A. T. Ihler, J. W. Fisher, R. L. Moses, A. S. Willsky, Nonparametric belief propagation for self-localization of sensor networks. IEEE Sel, J., Areas Commun. **23**(4), 809–819 (2005)
24. J. Lien, *A Framework for Cooperative Localization in Ultra-Wideband Wireless Networks*, (Cambridge, 2007)

25. F. Scheidt, D. Jin, M. Muma, A. M. Zoubir, in *Proc. 24th European Signal Process. Conf. (EUSIPCO)*. Fast and accurate cooperative localization in wireless sensor networks, (2016), pp. 190–194

26. D. Jin, F. Yin, C. Fritsche, A. M. Zoubir, F. Gustafsson, in *Proc. 23rd European Signal Process. Conf. (EUSIPCO)*. Efficient cooperative localization algorithm in LOS/NLOS environments, (2015), pp. 185–189

27. N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, R. J. O'Dea, Relative Location Estimation in Wireless Sensor Networks. IEEE Trans. Signal Process. **51**(8), 2137–2148 (2003)

28. R. W. Ouyang, K.-S. A.Wong, C.-T. Lea, Received Signal Strength-Based Wireless Localization via Semidefinite Programming: Noncooperative and Cooperative Schemes. IEEE Trans. Veh. Technol. **59**(3), 1307–1318 (2010)

29. H. Lim, L. C. Kung, J. C. Hou, H. Luo, in *Proc. 25th IEEE Int. Conf. Computer Commun. (INFOCOM)*. Zero-configuration, robust indoor localization: theory and experimentation, (Barcelona, 2006), pp. 1–12

30. A. Coluccia, F. Ricciato, in *Proc. IEEE Int. Sympos. Wireless Pervasive Comput. (ISWPC)*. On ML estimation for automatic RSS-based indoor localization, (Modena, 2010), pp. 495–502

31. J. Riba, A. Urruela, in *Proc. 29th IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP), vol. 2*. A non-line-of-sight mitigation technique based on ML-detection, (2004)

32. S. Gezici, Z. Sahinoglu, *UWB geolocation techniques for IEEE 802.15.4a personal area networks*, (Cambridge, 2004). TR-2004-110

33. J. J. Caffery, G. L. Stuber, Overview of radiolocation in CDMA cellular systems. IEEE Commun. Mag. **36**(4), 38–45 (1998)

34. S. Venkatesh, R. M. Buehrer, NLOS mitigation using linear programming in ultrawideband location-aware networks. IEEE Trans. Veh. Technol. **56**, 3182–3198 (2007)

35. X. Wang, Z. Wang, B. O'Dea, A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation. IEEE Trans. Veh. Technol. **52**(1), 112–116 (2003)

36. R. Casas, A. Marco, J. J. Guerrero, J. Falcó, Robust estimator for non-line-of-sight error mitigation in indoor localization. EURASIP Adv, J., Signal Process. **2006**(1), 043429 (2006)

37. G.-L.S., W.G., Bootstrapping M-estimators for reducing errors due to non-line-of-sight (NLOS) propagation. IEEE Commun. Lett. **8**(8), 509–510 (2004)

38. D. Liu, Y. Wang, P. He, Y. Zhai, H. Wang, TOA localization for multipath and NLOS environment with virtual stations. EURASIP Wireless, J., Commun. and Netw. **2017**(1), 104 (2017)

39. F. Yin, C. Fritsche, F. Gustafsson, A. M. Zoubir, TOA-based robust wireless geolocation and Cramér-Rao lower bound analysis in harsh LOS/NLOS environments. IEEE Trans. Signal Process. **61**(9), 2243–2255 (2013)

40. S. D. Chitte, S. Dasgupta, Z. Ding, Distance estimation from received signal strength under log-normal shadowing: bias and variance. IEEE Signal Process. Lett. **16**(3), 216–218 (2009)

41. A. Coluccia, Reduced-bias ML-based estimators with low complexity for self-calibrating RSS ranging. IEEE Trans. Wireless Commun. **12**(3), 1220–1230 (2013)

42. C. Savarese, J. M. Rabaey, J. Beutel, in *Proc. 26th IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*. Location in distributed ad-hoc wireless sensor networks, vol. 1, (2001), pp. 2037–2040

43. C. Savarese, J. M. Rabaey, K. Langendoen, in *Proc. Gen. Track Ann. Conf. USENIX Ann. Techn. Conf, ATEC '02*. Robust positioning algorithms for aistributed ad-hoc wireless sensor networks (USENIX Association, Berkeley, 2002), pp. 317–327

44. A. Savvides, C.-C. Han, in *Proc. 7th Ann. Int. Conf. Mobile Comput. Netw. (MobiCom). MobiCom '01*. M.B.S.: Dynamic fine-grained localization in ad-hoc networks of sensors (ACM, New York, 2001), pp. 166–179

45. S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. (Prentice Hall, Englewood Cliffs, New Jersey, 1993)

46. L. Mihaylova, D. Angelova, D. Bull, N. Canagarajah, Localization of mobile nodes in wireless networks with correlated in time measurement noise. IEEE Trans. Mobile Comput. **10**(1), 44–53 (2011)

47. B. Beck, S. Lanh, R. Baxley, X. Ma, Uncooperative emitter localization using signal strength in uncalibrated mobile networks. IEEE Trans. Wireless Commun. **16**(11), 7488–7500 (2017)