# Blockage Intelligence in Complex Environments for Beyond 5G Localization

Gianluca Torsoli, *Student Member, IEEE*, Moe Z. Win, *Fellow, IEEE*, and Andrea Conti, *Fellow, IEEE*

***Abstract*— Location awareness is vital for several applications in wireless ecosystems, including fifth generation (5G) and beyond networks defined by the 3rd Generation Partnership Project (3GPP). However, complex wireless environments such as indoor factories are characterized by harsh multipath propagation and non-line-of-sight (NLOS) conditions, which are detrimental to localization accuracy. This paper introduces the concept of blockage intelligence (BI) to provide a probabilistic description of wireless propagation conditions. Then, it discusses its integration in both conventional and soft information (SI)-based localization algorithms. Case studies are presented in the 3GPP indoor factory scenario with various gNodeBs (gNBs) deployments. Results show that BI together with SI-based localization significantly outperforms existing localization techniques. The rich information provided by BI is vital to perform accurate localization in 5G and beyond networks operating in complex wireless environments.**

***Index Terms*— 5G, localization, NLOS identification, wireless networks, industrial IoT.**

## I. Introduction

LOCATION AWARENESS is a key enabler for a myriad of applications in fifth generation (5G) and beyond wireless networks [1], [2], [3], [4], [5], including autonomy [6], [7], [8], [9], [10], smart environments [11], [12], [13], [14], [15], assets tracking [16], [17], [18], [19], [20], and Internet-of-Things (IoT) [21], [22], [23], [24]. In particular, with the rapid transition towards Industry 4.0, the interest in accurate localization for Industrial IoT (IIoT) applications is increasing rapidly [25], [26], [27], [28]. The localization requirements in 5G networks for the different applications are defined by 3rd Generation Partnership Project (3GPP) in terms of key performance indicators (KPIs) which include accuracy, latency, and availability [29]. However, fulfilling such requirements in complex wireless environments is particularly challenging. Industrial sites are complex wireless environments, typically characterized by a large number of metallic surfaces and machines, which determine heavy multipath propagation and frequent non-line-of-sight (NLOS) conditions [30]. These impairments are particularly detrimental to localization accuracy, especially at millimeter waves [31], [32], [33].

According to current 3GPP specifications, localization in 5G networks relies on processing single-value estimates (SVEs) extracted from the received waveforms such as time difference-of-arrival (TDOA), round-trip time (RTT), and angle-of-departure (AOD) through conventional SVE-based algorithms. To improve localization performance, also information on additional paths for multipath mitigation and NLOS identification can be employed [34], [35], [36], [37], [38]. However, SVE-based localization algorithms are typically unable to satisfy the KPI requirements for IIoT applications due to the insufficient positional information encapsulated in SVEs [39]. In this context, 3GPP study items for Release 18 are investigating the use of machine learning (ML) and artificial intelligence to improve 5G networks localization and communication capabilities [40]. The support to ML solutions is fundamental to enable more effective algorithms for localization in 5G and beyond wireless networks [41], [42]. In particular, soft information (SI)-based localization can effectively leverage ML techniques to overcome the limitations of conventional SVE-based algorithms [43], [44]. Specifically, SI-based localization makes use of probabilistic models to characterize the relationship between measurements, contextual information, and user equipment (UE) position in 5G and beyond networks [5]. However, the extraction of SI from SVEs provided by 5G measurements alone may be insufficient to provide accurate localization in complex wireless environments, such as industrial scenarios. To address this issue, NLOS identification can be seamlessly integrated in SI-based localization [43]. However, conventional NLOS identification techniques provide binary information that is not able to provide rich information on the wireless propagation conditions [37], [45], [46]. Moreover, the prior geometrical or statistical characterization of the wireless environments typically required by such methods is not always available [38], [47], [48]. Therefore, it is necessary to develop more effective NLOS identification techniques, able to provide richer and more flexible information.

Fundamental questions related to NLOS identification for 5G networks in complex wireless environments are:

- how the positional information encapsulated in the 5G received signals can be leveraged to enhance NLOS identification; and

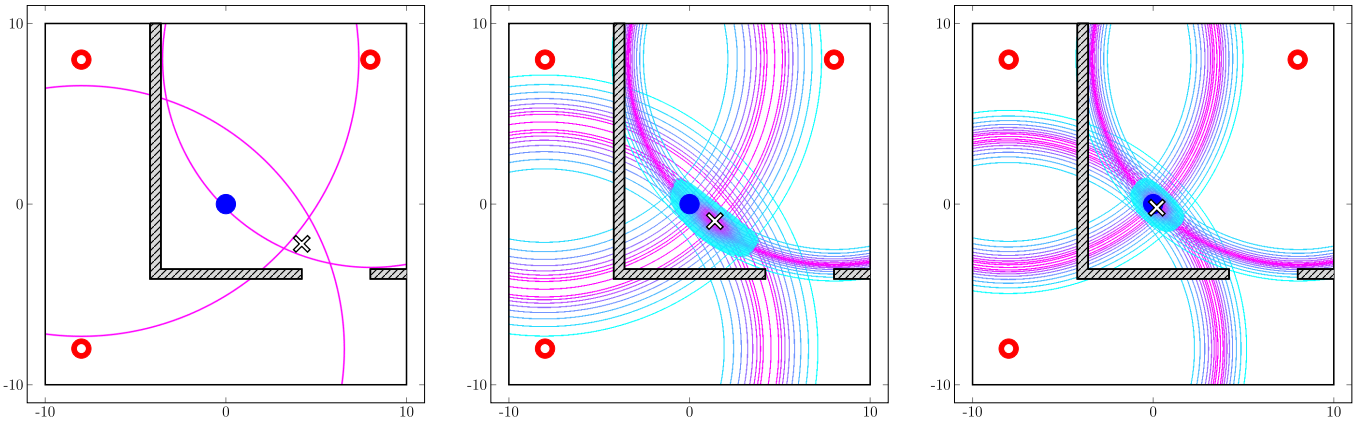
Manuscript received 25 October 2022; revised 24 April 2023; accepted 28 April 2023. Date of publication 12 May 2023; date of current version 19 June 2023. The fundamental research described in this paper was supported in part by the Office of Naval Research under Grant N62909-22-1-2009, in part by the National Science Foundation under Grant CNS-2148251, and in part by the Federal Agency and Industry Partners in the RINGS Program. The material in this paper was presented, in part, at the IEEE/ION Position Location and Navigation Symposium, Monterey, CA, USA, April 2023. *(Corresponding author: Andrea Conti.)*

Gianluca Torsoli and Andrea Conti are with the Department of Engineering and CNIT, University of Ferrara, 44122 Ferrara, Italy (e-mail: gianluca.torsoli@unife.it; a.conti@ieee.org).

Moe Z. Win is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: moewin@mit.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/JSAC.2023.3275612.

Digital Object Identifier 10.1109/JSAC.2023.3275612



(a) SVE-based localization (b) SI-based localization (c) SI-based localization with BI

Fig. 1. Pictorial view of SVE-based localization, SI-based localization and SI-based localization with BI. The red annulus indicates a gNB, the blue circle indicates the real UE position and the cross indicates the estimated position. The hatched polygons represent obstacles that determine NLOS conditions. The coordinates on the axis are in meters.

- what is the performance gain that a probabilistic characterization of NLOS conditions can provide to localization algorithms?

The answers to these questions will enable the development of localization techniques able to fulfill service-level requirements for 5G and beyond wireless networks operating in complex wireless environments. The goal of this paper is to provide a probabilistic representation of wireless propagation conditions, referred to as blockage intelligence (BI), that can be integrated into both SVE-based and SI-based localization. On the one hand, BI can be used to weigh the quality of measurements in SVE-based localization. On the other hand, the probabilistic information provided by BI enables to reduce the uncertainty in SI models, thus improving the localization accuracy as depicted in Fig. 1. The key idea of BI is to process the rich information encapsulated in the received signals to provide a probabilistic characterization of NLOS propagation conditions. Such approach is particularly suitable for 5G and beyond wireless networks given the support to soft NLOS identification values starting from 3GPP Release 17 [49]. Accordingly, we advocate that the integration of BI in the SI framework is able to provide a new level of accuracy for localization in 5G and beyond complex wireless environments.

This paper proposes the concept of BI to enhance location awareness of 5G and beyond networks operating in complex wireless environments. The key contributions of the paper can be summarized as follows:

- proposal of BI for providing a probabilistic characterization of NLOS propagation conditions;
- development of a low-complexity method for obtaining BI in complex wireless conditions; and
- quantification of the localization performance gain provided by BI in the 3GPP indoor factory (InF) scenario.

The remainder of the paper is organized as follows. Section II briefly describes localization in 5G networks. Section III introduces a method for obtaining BI and discusses its integration in localization algorithms. Section IV presents two case studies in 3GPP-standardized InF scenarios. Finally, Section V gives our conclusions.

*Notations*: Random variables are displayed in sans serif, upright fonts; their realizations in serif, italic fonts. Vectors and matrices are denoted by bold lowercase and uppercase letters, respectively. For example, a random variable and its realization are denoted by $\mathsf{x}$ and $x$, respectively; a random vector and its realization are denoted by $\mathbf{x}$ and $\boldsymbol{x}$, respectively; a matrix is denoted by $\boldsymbol{X}$. For a matrix $\boldsymbol{X}$ and a vector $\boldsymbol{x}$, their transpose is denoted by $\boldsymbol{X}^{\mathrm{T}}$ and $\boldsymbol{x}^{\mathrm{T}}$, respectively. $\|\boldsymbol{x}\|$ denotes the norm of the vector $\boldsymbol{x}$. Sets are denoted by calligraphic font. For example, a set is denoted by $\mathcal{X}$. $x^*$ denotes the complex conjugate of $x$; $|x|$ denotes the absolute value of $x$; $\lceil x \rceil$ denotes the smallest integer greater or equal to $x$; The function $f_{\mathbf{x}}(\boldsymbol{x})$ and, for brevity when possible, $f(\boldsymbol{x})$ denotes the probability density function (PDF) of a continuous random vector $\mathbf{x}$; $\varphi(\boldsymbol{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the PDF of a Gaussian random vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. $\mathbb{E}_{\mathsf{x}|\mathsf{y}}\{\cdot|\boldsymbol{y}\}$ denotes the expectation with respect to the random variable $\mathsf{x}$ conditional on $\mathbf{y} = \boldsymbol{y}$. The function $\mathbb{1}_{\mathcal{A}}(x)$ is equal to 1 if $x \in \mathcal{A}$ and 0 otherwise. The notation $O(\cdot)$ is used to indicate asymptotic complexity.

## II. Localization in 5G Networks

Consider a 5G localization network composed of $N_{\mathrm{g}}$ base stations, namely gNodeBs (gNBs), indexed by $j \in \mathcal{N}_{\mathrm{g}} = \{1, 2, \ldots, N_{\mathrm{g}}\}$ and with known positions $\{\boldsymbol{p}_j\}_{j\in\mathcal{N}_{\mathrm{g}}}$. Based on a collection of measurements $\{\boldsymbol{y}_j\}_{j\in\mathcal{N}_{\mathrm{g}}}$ obtained from the gNBs, it is possible to estimate the UE position. Measurements may include time-, angle-, or power-based metrics, waveform samples, and any combination of them. Each measurement $\boldsymbol{y}_j$ is related to a positional feature $\theta_j(\boldsymbol{p})$ which depends on both the UE and the gNB positions.[1] According to current 3GPP specifications, $\{\boldsymbol{y}_j\}_{j\in\mathcal{N}_{\mathrm{g}}}$ must be processed to obtain a set of SVEs of positional features, denoted by $\{\hat{\theta}_j\}_{j\in\mathcal{N}_{\mathrm{g}}}$, which can be used for localization [49].

### A. SVEs for 5G Localization

3GPP standardization documents define two reference signals (RSs), namely the positioning reference signal (PRS) and the sounding reference signal (SRS) for downlink (DL)

[1]In the following, the dependence on $\boldsymbol{p}$ of measurements and features will be omitted for notational brevity.

and uplink (UL) localization, respectively. While DL localization can rely on the transmission of a PRS from a gNB to an UE, UL localization can rely on the transmission of an SRS from an UE to a gNB [50]. PRS and SRS share a similar structure and are generated as orthogonal frequency division multiplexing (OFDM) signals. The PRS and SRS can be transmitted with different carrier frequencies either in frequency range 1 (FR1) (carrier frequency between 410 MHz and 7.125 GHz) or frequency range 2 (FR2) (carrier frequency between 24.25 GHz and 52.6 GHz). RSs transmission can be performed with different numerologies (i.e., different bandwidth and subcarrier spacings) and with different time-frequency configurations [51]. By processing the received RSs, one can extract SVEs for localization, including time-of-arrival (TOA), TDOA, RTT, and AOD [50], [52].

TOA is typically determined based on the estimation of the delay associated with the first path in the cross-correlation between the transmitted and the received RS. However, in complex wireless environments, NLOS conditions and multipath propagation introduce biases in TOA estimates that are detrimental to localization accuracy. Different approaches have been explored to mitigate the biases in TOA estimates [38], [53], [54]. We consider the iterative algorithm proposed in [55] for TOA estimation, which is briefly reported in the following. Let $r(t)$ and $s(t)$ be respectively the received and the transmitted RS (either PRS or SRS) with samples $r[n] = r(nT_\mathrm{s})$ and $s[n] = s(nT_\mathrm{s})$, where $T_\mathrm{s}$ denotes the sampling time. Let $R_{r,s}[n]$ denote the cross-correlation between $r[n]$ and $s[n]$, which is given by

$$R_{r,s}[n] = \sum_{k=0}^{N_\mathrm{s}-1} r[k]\, s^*[n-k] \tag{1}$$

for $n = 0, 1, \ldots, N_\mathrm{s} - 1$ where $N_\mathrm{s}$ denotes the number of samples per waveform. At each iteration, the sample associated with the peak of $R_{r,s}^{[i]}$ is identified, and then its contribution is removed from the cross-correlation considered for the next iteration. After $N_\mathrm{I}$ iterations the TOA is estimated as the smallest delay (which corresponds to the first path) detected during the iterative procedure. The use of TOA measurements for localization requires accurate synchronization between the UEs and the gNBs. This represents the main challenge for the adoption of TOA-based localization and, starting from Release 17, the 3GPP put forth methods for time synchronization of all the network elements, including the UEs [56]. In particular, time synchronization is widely considered for 5G and beyond dedicated wireless networks, especially in industrial environments [57], [58], [59], [60].

From TOA measurements, RTT measurements are obtained averaging DL-TOA (obtained from PRS transmission) and UL-TOA (obtained from SRS transmission). The estimation of the AOD can be performed if the gNBs are equipped with a sufficient number of antennas to enable beamforming. Specifically, from multiple PRS transmissions with different beam orientation angles, the AOD can be estimated by identifying the beam with the highest reference signal received power (RSRP) [61].

### B. SVE-Based Localization

Conventional SVE-based localization algorithms directly exploit the SVEs to perform localization. For example, the least squares (LS) algorithm is widely considered for estimating the UE position in mobile networks [62], [63], [64]. Let $\{\hat{\theta}_j\}_{j\in\mathcal{N}_\mathrm{g}}$ be a set of SVEs obtained from the gNBs, then the UE position is given by

$$\hat{\boldsymbol{p}} = \arg\min_{\tilde{\boldsymbol{p}}} \left\| \hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\tilde{\boldsymbol{p}}) \right\|^2 \tag{2}$$

where $\|\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\tilde{\boldsymbol{p}})\|$ must be intended in terms of angular distance for localization with angle-based SVEs and

$$\hat{\boldsymbol{\theta}} = \left[\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_{N_\mathrm{g}}\right]^\mathrm{T} \tag{3a}$$

$$\boldsymbol{\theta}(\tilde{\boldsymbol{p}}) = \left[\theta_1(\tilde{\boldsymbol{p}}), \theta_2(\tilde{\boldsymbol{p}}), \ldots, \theta_{N_\mathrm{g}}(\tilde{\boldsymbol{p}})\right]^\mathrm{T}. \tag{3b}$$

Note that the use of SVEs only is a limitation to the localization accuracy since SVEs are typically unable to fully provide the positional information encapsulated in the waveform.

### C. SI-Based Localization

SI-based localization leverages ML techniques to provide a probabilistic characterization of the relationship between UE position, measurements, and environmental information [43]. In particular, this is performed by considering both soft feature information (SFI) (i.e., the positional information encapsulated by the measurements), and soft context information (SCI) (i.e., the positional information related to contextual data). In a non-Bayesian setting, the SFI related to the SVE $\hat{\theta}$ is a function of the positional feature as

$$\mathcal{L}_{\hat{\theta}}(\theta) \propto f_{\hat{\theta}}(\hat{\theta}; \theta)\,. \tag{4}$$

Given a collection of independent SVEs $\{\hat{\theta}_j\}_{j\in\mathcal{N}_\mathrm{g}}$ obtained from different gNBs, a maximum likelihood estimation can be performed to infer the UE position as[2]

$$\hat{\boldsymbol{p}} = \arg\max_{\tilde{\boldsymbol{p}}} \prod_{j\in\mathcal{N}_\mathrm{g}} \mathcal{L}_{\hat{\theta}_j}(\theta_j(\tilde{\boldsymbol{p}}))\,. \tag{5}$$

In addition, SI-based localization enables data-fusion. Specifically, if each gNB is able to provide conditionally independent heterogeneous measurements, their SFIs can be efficiently fused via multiplication [43]. SFI is estimated as proportional to the joint probability distributions of measurements and positional features, hereafter referred to as the generative model. In complex wireless environments, the generative models cannot be determined a priori and it is necessary to learn them from the environment via density estimation. We consider generative models in the class of Gaussian mixture models (GMMs) which have been recently demonstrated to be effective in 5G and beyond wireless networks [44]. Let $\boldsymbol{x} = \left[\hat{\theta}\,,\, \theta\right]^\mathrm{T}$ and $\boldsymbol{z} = \mathcal{S}(\boldsymbol{x})$ where $\mathcal{S}$ denotes a data sphering transformation [65]. Then, the generative model approximation in terms of GMM is given by

$$f(\boldsymbol{z}; \Xi) = \sum_{i=1}^{N_\mathrm{m}} \alpha_i \varphi(\boldsymbol{z}; \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i) \tag{6}$$

[2]If also SCI is considered, (5) becomes a maximum a posteriori estimation.

where $\Xi = \{\alpha_i, \boldsymbol{\mu}_i, \boldsymbol{\Sigma}_i\}_{i=1}^{N_\mathrm{m}}$ is the set of parameters of the GMM, $N_\mathrm{m}$ is the number of Gaussian components, and $\alpha_i \in \mathbb{R}^+$ are the mixing coefficients such that $\sum_{i=1}^{N_\mathrm{m}} \alpha_i = 1$. The use of GMM enables the approximation of complex densities if $N_\mathrm{m}$ and $\Xi$ are accurately estimated. To this end, given a sphered training dataset $\{\boldsymbol{z}_n\}_{n=1}^{N_\mathrm{D}}$, an estimate $\hat{\Xi}$ of the set of parameters can be determined through maximum likelihood estimation. Specifically, since a closed-form solution cannot be obtained, $\hat{\Xi}$ can be approximated through the expectation-maximization algorithm [66]. Details about the importance of the dataset dimension for training the SI generative model can be found in [5].

## III. Blockage Intelligence

Localization algorithms benefit from the positional information provided by NLOS identification. However, the lack of a probabilistic characterization of NLOS conditions and the need for prior geometrical knowledge of the environment limit the potential of NLOS identification in localization. To overcome such limitations, we exploit the intrinsic information on wireless propagation encapsulated in the received signals. This can be obtained by processing the cross-correlation between the transmitted and the received RSs. On the one hand, such cross-correlation captures rich information on the wireless channel conditions; on the other, it is commonly available and necessary to many ranging algorithms for TOA estimation [35]. In particular, we consider the positional information encapsulated in

$$g[n] = \left|R_{r,s}[n]\right| \tag{7}$$

for $n = 0, 1, \ldots, N_\mathrm{c} - 1$ where $N_\mathrm{c} = \lceil T_\mathrm{M}/T_\mathrm{s} \rceil$ and $T_\mathrm{M}$ is the maximum path delay such that the received waveform contains positional information (e.g., $T_\mathrm{M}$ can be determined based on the maximum distance between the gNB and the farthest obstacle or based on the wireless channel delay spread). Fig. 2 shows an example of realization of $g[n]$ in line-of-sight (LOS) and NLOS conditions. It can be observed that in LOS conditions, $g[n]$ exhibits a dominant peak around the real TOA between the gNB and the UE. Conversely, $g[n]$ in NLOS conditions is characterized by several peaks with similar amplitudes spread over a wide time frame. Note that in such case the highest peak in $g[n]$ is due to a signal replica caused by multipath propagation.

### A. Feature Extraction

Despite the rich information contained in $g[n]$, its use can be unfeasible for several applications due to its high dimensionality, especially when the RSs are transmitted with wide bandwidths. Therefore, a dimensionality reduction step on $g[n]$ is needed to effectively implement BI. In particular, a low-dimensional set of features can be computed to extract the positional information encapsulated by $g[n]$. Such features can be selected based on the time and amplitude dispersions that $g[n]$ exhibits in different propagation conditions [37], [38], [67], [68]. These characteristics can be quantified by computing statistical indicators on $g[n]$. Differently from TOA estimation, the computation of such statistical indicators for BI

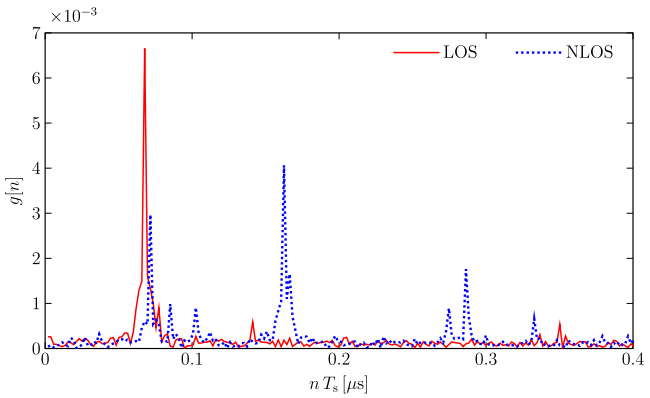



Fig. 2. Example of $g[n]$ in LOS and NLOS conditions considering a UE in a fixed position distant $d = 20\,\mathrm{m}$ from the gNB and $T_\mathrm{M} = 0.4\,\mu\mathrm{s}$ with perfect syncronization between the gNB and the UE, the real TOA is $d/c \simeq 0.067\,\mu\mathrm{s}$, where $c$ denotes the speed of light.

does not require accurate syncronization between the UE and the gNBs. Specifically, we evaluate the time dispersion of $g[n]$ considering its normalized version given by

$$\breve{g}[n] = \frac{g[n]}{\sum_{m=0}^{N_\mathrm{c}-1} g[m]}. \tag{8}$$

Then, it is possible to compute the time dispersion mean, variance, skewness, and kurtosis, which are respectively given by

$$\mu_\mathrm{t} = \sum_{m=0}^{N_\mathrm{c}-1} mT_\mathrm{s}\, \breve{g}[m] \tag{9a}$$

$$\sigma_\mathrm{t}^2 = \sum_{m=0}^{N_\mathrm{c}-1} \left(mT_\mathrm{s} - \mu_\mathrm{t}\right)^2 \breve{g}[m] \tag{9b}$$

$$\chi_\mathrm{t} = \frac{\sum_{m=0}^{N_\mathrm{c}-1} \left(mT_\mathrm{s} - \mu_\mathrm{t}\right)^3 \breve{g}[m]}{\left(\sigma_\mathrm{t}^2\right)^{\frac{3}{2}}} \tag{9c}$$

$$\kappa_\mathrm{t} = \frac{\sum_{m=0}^{N_\mathrm{c}-1} \left(mT_\mathrm{s} - \mu_\mathrm{t}\right)^4 \breve{g}[m]}{\left(\sigma_\mathrm{t}^2\right)^2}. \tag{9d}$$

To evaluate the amplitude dispersion, we consider information extracted directly from $g[n]$, namely, energy, maximum value, sample mean, sample variance, sample skewness, and sample kurtosis which are respectively given by

$$\mathcal{E} = \sum_{m=0}^{N_\mathrm{c}-1} g[m]^2 \tag{10a}$$

$$M = \max_m g[m] \tag{10b}$$

$$\mu_\mathrm{a} = \frac{1}{N_\mathrm{c}} \sum_{m=0}^{N_\mathrm{c}-1} g[m] \tag{10c}$$

$$\sigma_\mathrm{a}^2 = \frac{1}{N_\mathrm{c}} \sum_{m=0}^{N_\mathrm{c}-1} \left(g[m] - \mu_\mathrm{a}\right)^2 \tag{10d}$$

$$\chi_\mathrm{a} = \frac{1}{N_\mathrm{c}} \frac{\sum_{m=0}^{N_\mathrm{c}-1} \left(g[m] - \mu_\mathrm{a}\right)^3}{\left(\sigma_\mathrm{a}^2\right)^{\frac{3}{2}}} \tag{10e}$$

$$\kappa_\mathrm{a} = \frac{1}{N_\mathrm{c}} \frac{\sum_{m=0}^{N_\mathrm{c}-1} \left(g[m] - \mu_\mathrm{a}\right)^4}{\left(\sigma_\mathrm{a}^2\right)^2}. \tag{10f}$$

The vector of statistical features containing the positional information exploited by BI is

$$\boldsymbol{\nu} = \left[\, \mu_{\mathrm{t}},\, \sigma^2_{\mathrm{t}},\, \chi_{\mathrm{t}},\, \kappa_{\mathrm{t}},\, \mathcal{E},\, M,\, \mu_{\mathrm{a}},\, \sigma^2_{\mathrm{a}},\, \chi_{\mathrm{a}},\, \kappa_{\mathrm{a}} \,\right]^{\mathrm{T}}. \tag{11}$$

### B. BI Probabilistic Model

In this section, the statistical indicators defined in Section III-A are employed to develop a method for obtaining BI. The use of a probabilistic characterization of NLOS propagation conditions enables overcoming the limitation of binary NLOS identification. In particular, a probabilistic characterization of NLOS propagation conditions can be obtained considering a two-class supervised classification problem [69]. Let $\Upsilon \in \mathcal{C} = \{+1, -1\}$ be a binary random variable which takes value $+1$ and $-1$ for NLOS and LOS propagation conditions, respectively, and let $\mathbf{v} \in \mathbb{R}^d$ be a $d$-dimensional random vector of features. By considering an exponential loss function, a model $c(\boldsymbol{\nu})$ for classification can be obtained as

$$c(\boldsymbol{\nu}) = \underset{\check{c}\,:\,\mathbb{R}^d \to \mathbb{R}}{\arg\min}\, \mathbb{E}_{\Upsilon|\mathbf{v}}\{e^{-\Upsilon \check{c}(\mathbf{v})} \mid \boldsymbol{\nu}\}. \tag{12}$$

By observing that

$$\begin{aligned}\mathbb{E}_{\Upsilon|\mathbf{v}}\{e^{-\Upsilon c(\mathbf{v})} \mid \boldsymbol{\nu}\} &= \mathbb{P}\{\Upsilon = +1 \mid \boldsymbol{\nu}\}\, e^{-c(\boldsymbol{\nu})} \\ &\quad + (1 - \mathbb{P}\{\Upsilon = +1 \mid \boldsymbol{\nu}\})\, e^{c(\boldsymbol{\nu})} \qquad (13)\\ \frac{\partial \mathbb{E}_{\Upsilon|\mathbf{v}}\{e^{-\Upsilon c(\mathbf{v})} \mid \boldsymbol{\nu}\}}{\partial c(\boldsymbol{\nu})} &= -\mathbb{P}\{\Upsilon = +1 \mid \boldsymbol{\nu}\}\, e^{-c(\boldsymbol{\nu})} \\ &\quad + (1 - \mathbb{P}\{\Upsilon = +1 \mid \boldsymbol{\nu}\})\, e^{c(\boldsymbol{\nu})} \qquad (14)\end{aligned}$$

and by setting (14) equal to zero, a closed-form solution for (12) can be obtained as [70]

$$c(\boldsymbol{\nu}) = \frac{1}{2} \log\left(\frac{\mathbb{P}\{\Upsilon = +1 \mid \boldsymbol{\nu}\}}{1 - \mathbb{P}\{\Upsilon = +1 \mid \boldsymbol{\nu}\}}\right). \tag{15}$$

Hence, the probability of NLOS conditions given a set of statistical features of the received signal results in

$$\psi(\boldsymbol{\nu}) = \mathbb{P}\{\Upsilon = +1 \mid \boldsymbol{\nu}\} = \frac{e^{c(\boldsymbol{\nu})}}{e^{-c(\boldsymbol{\nu})} + e^{c(\boldsymbol{\nu})}}. \tag{16}$$

However, solving (12) requires the joint probability distribution of $\Upsilon$ and $\mathbf{v}$, which is not known a priori. Therefore, it is necessary to approximate $c(\boldsymbol{\nu})$ via empirical risk minimization (ERM) leveraging a training dataset obtained through a measurement campaign. In particular, the choice of an exponential loss function in (12) enables the use of the Real Adaboost algorithm for ERM [70]. Such algorithm aims to fit a set of $N_{\mathrm{L}}$ weak learners (i.e., classifiers that perform slightly better than a random guess) to approximate $c(\boldsymbol{\nu})$ via additive modeling [71]. This is typically obtained using weak learners with high bias and low variance such as one-level classification and regression trees (CARTs). The parameters of the weak learners are fit iteratively, giving at each iteration an increased weight to the training samples that were misclassified by the weak learner fit at the previous iteration [72]. Specifically, let $\mathcal{D} = \left\{(\boldsymbol{\nu}_n, \gamma_n)\right\}_{n=1}^{N_{\mathrm{D}}}$ be a training dataset of $N_{\mathrm{D}}$ realizations of $(\mathbf{v}, \Upsilon)$. Then, the offline (training) phase of the proposed method for BI proceeds as described in Algorithm 1 [70], [73].

**Algorithm 1** BI Model Training

1. Initialize the weights $w_n^{[0]} = 1/N_{\mathrm{D}}$ for $n = 1, 2, \ldots, N_{\mathrm{D}}$.
2. **for** $m = 1, 2, \ldots, N_{\mathrm{L}}$ **do**
   3. Fit a weak learner $T^{[m]}(\boldsymbol{\nu})$ to the training data using the weights $w_n^{[m]}$.
   4. Set the basis function as
   $$\beta^{[m]}(\boldsymbol{\nu}) = \frac{1}{2} \log\left(\frac{\hat{\pi}^{[m]}(\boldsymbol{\nu})}{1 - \hat{\pi}^{[m]}(\boldsymbol{\nu})}\right) \tag{17}$$
   where $\hat{\pi}^{[m]}(\boldsymbol{\nu})$ is a NLOS probability estimate obtained from $T^{[m]}(\boldsymbol{\nu})$.
   5. Update the training samples weights for the next iteration as
   $$w_n^{[m+1]} = w_n^{[m]} e^{-\gamma_n \beta^{[m]}(\boldsymbol{\nu}_n)}. \tag{18}$$
   6. Normalize the weights such that $\sum_{n=1}^{N_{\mathrm{D}}} w_n^{[m+1]} = 1$.
7. **end for**

In the online phase, the weak learners are combined to obtain an approximation $\tilde{c}(\boldsymbol{\nu})$ of $c(\boldsymbol{\nu})$ which is given by

$$\tilde{c}(\boldsymbol{\nu}) = \sum_{m=1}^{N_{\mathrm{L}}} \beta^{[m]}(\boldsymbol{\nu}) \tag{19}$$

where $\beta^{[m]}(\boldsymbol{\nu})$ is the basis function obtained from the $m$-th weak learner.[3]

The probabilistic information provided by BI can reduce to conventional binary NLOS identification, hereafter referred to as discretized blockage intelligence (DBI), by simply applying a threshold to $\psi(\boldsymbol{\nu})$, i.e.,

$$\hat{\psi} = \begin{cases} +1, & \psi(\boldsymbol{\nu}) \geqslant 0.5 \\ -1, & \psi(\boldsymbol{\nu}) < 0.5 \end{cases} \tag{20}$$

thus resulting in binary NLOS identification as a special case.

*Remark:* The proposed framework for BI does not rely on any specific characteristic of $\boldsymbol{\nu}$. Hence, any quantity which encapsulates information on the wireless propagation conditions can be exploited by BI. Such information can be both provided by the 5G network or obtained from different sources (e.g., inertial sensors, Bluetooth, and Wi-Fi).

### C. Classification and Regression Trees

One-level CARTs are commonly employed as weak learners in the Real Adaboost algorithm [70]. In particular, they split the multidimensional feature space into two regions, namely, $\mathcal{R}_{+1}^{(\hat{\boldsymbol{\varphi}})}$ and $\mathcal{R}_{-1}^{(\hat{\boldsymbol{\varphi}})}$, which represent the NLOS and LOS classification regions, respectively. The vector $\hat{\boldsymbol{\varphi}} = [\hat{q}, \hat{\eta}]$ contains the parameters that define the regions, i.e., the index $\hat{q} \in \mathcal{Q} = \{1, 2, \ldots, d\}$ of the single feature exploited by the one-level CART in the splitting condition and the respective threshold $\hat{\eta} \in \mathbb{R}$. Formally, the two regions are given by

$$\mathcal{R}_{+1}^{(\hat{\boldsymbol{\varphi}})} = \left\{\boldsymbol{\nu} \in \mathbb{R}^d : \nu_{\hat{q}} \geqslant \hat{\eta}\right\} \tag{21a}$$

[3]Note that (19) provides an approximation of $c(\boldsymbol{\nu})$ up to a scale factor. However, this scale factor does not impact the probabilistic characterization of NLOS conditions provided $\psi(\boldsymbol{\nu})$ in (16).

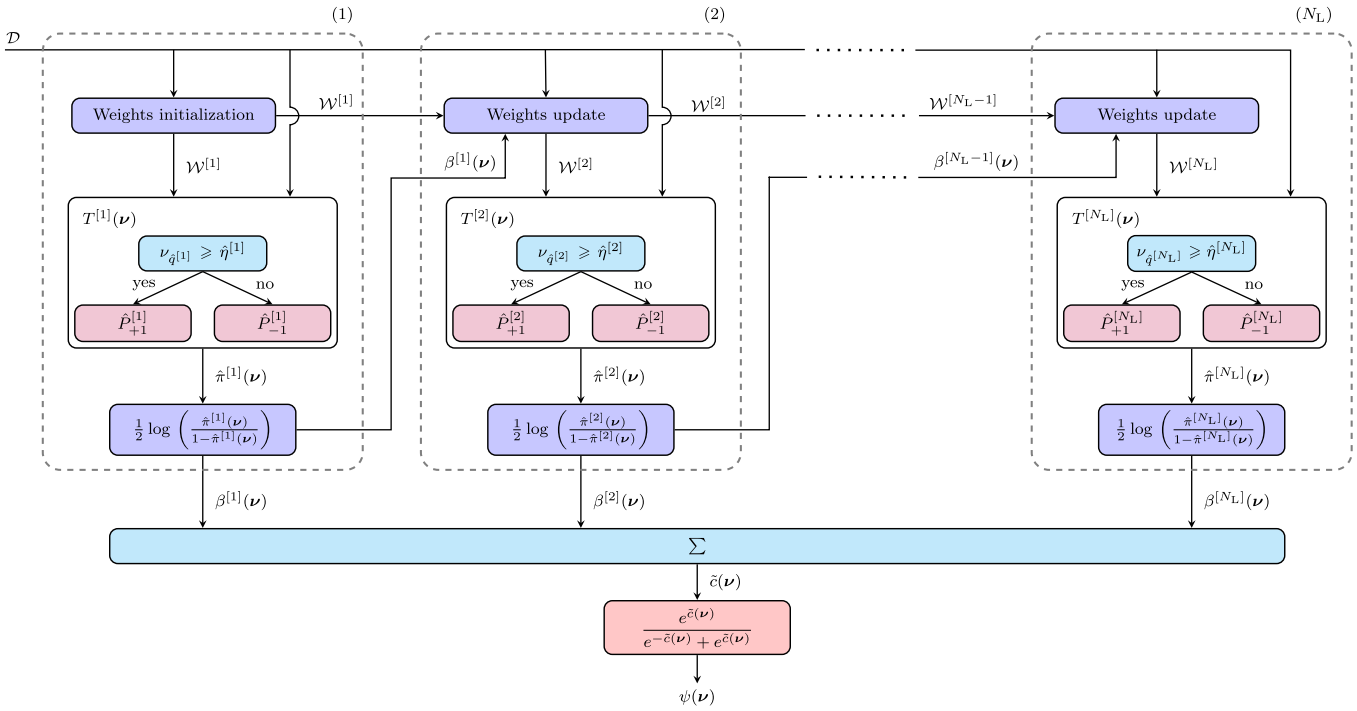



Fig. 3. Block diagram of the training phase of the proposed method for BI. The set $\mathcal{W}^{[m]} = \{w_n^{[m]}\}_{n=1}^{N_{\mathrm{D}}}$ denotes the weights exploited at the $m$-th iteration.

$$\mathcal{R}_{-1}^{(\hat{\boldsymbol{\varphi}})} = \left\{\boldsymbol{\nu} \in \mathbb{R}^d : \nu_{\hat{q}} < \hat{\eta}\right\}. \tag{21b}$$

Let $\mathcal{D} = \left\{(\boldsymbol{\nu}_n, \gamma_n)\right\}_{n=1}^{N_{\mathrm{D}}}$ be a training dataset, and let $\hat{P}_{+1}$ and $\hat{P}_{-1}$ denote the proportion of the training samples classified as belonging to $\mathcal{R}_{+1}^{(\boldsymbol{\varphi})}$ and $\mathcal{R}_{-1}^{(\boldsymbol{\varphi})}$, respectively, where $\hat{P}_{+1} = 1 - \hat{P}_{-1}$ [71]. Then, in an offline (training) phase, $\hat{q}$ and $\hat{\eta}$ are obtained by solving the following optimization problem

$$\hat{\boldsymbol{\varphi}} = \arg\min_{\boldsymbol{\varphi}} L(\boldsymbol{\varphi}) \tag{22}$$

where

$$L(\boldsymbol{\varphi}) = \sum_{c \in \mathcal{C}} \hat{P}_c \, G_c(\boldsymbol{\varphi}) \tag{23a}$$

$$G_c(\boldsymbol{\varphi}) = \sum_{c' \in \mathcal{C}} \breve{P}_{cc'}(\boldsymbol{\varphi})(1 - \breve{P}_{cc'}(\boldsymbol{\varphi})) \tag{23b}$$

$$\breve{P}_{cc'}(\boldsymbol{\varphi}) = \frac{\sum_{\boldsymbol{\nu}_n \in \mathcal{R}_c^{(\boldsymbol{\varphi})}} w_n \, \mathbb{1}_{\{c'\}}(\gamma_n)}{\sum_{\boldsymbol{\nu}_n \in \mathcal{R}_c^{(\boldsymbol{\varphi})}} w_n} \tag{23c}$$

and $w_n$ is the weight assigned to the $n$-th training sample. Equation (23b) is referred to as weighted Gini impurity measure and it is used as a cost function for (22) [71]. To solve the optimization problem in (22), an approach based on exhaustive search is typically feasible due to the simple form of (23a) [71]. Let $\mathcal{V}_q$ denote the set of all the values of the $q$-th feature among all the $\boldsymbol{\nu}_n \in \mathcal{D}$. For each $q \in \mathcal{Q}$ and for each $\eta \in \mathcal{V}_q$, (23a) is evaluated, and the couple $\left[\hat{q}, \hat{\eta}\right]$ which provides the minimum $L(\boldsymbol{\varphi})$ is selected to define the one-level CART splitting condition.

In the online phase, CARTs provide NLOS probability estimates as

$$\hat{\pi}(\boldsymbol{\nu}) = \hat{P}_{+1} \, \mathbb{1}_{\mathcal{R}_{+1}^{(\hat{\boldsymbol{\varphi}})}}(\boldsymbol{\nu}) + \hat{P}_{-1} \, \mathbb{1}_{\mathcal{R}_{-1}^{(\hat{\boldsymbol{\varphi}})}}(\boldsymbol{\nu}). \tag{24}$$

Fig. 3 depicts a complete flowchart of the training phase of the proposed method for BI, comprising the use of CARTs as weak learners.

### D. BI Complexity

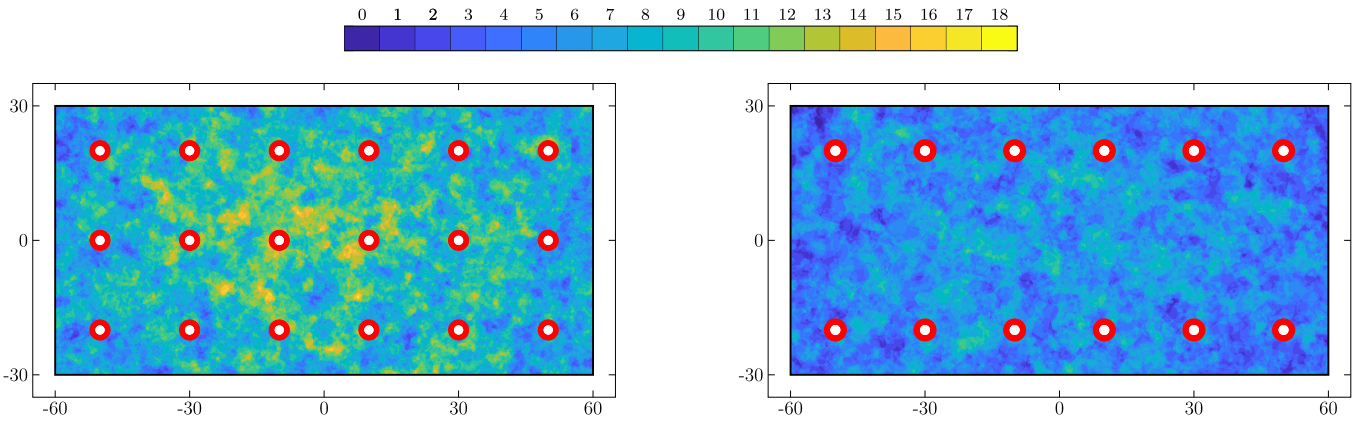
Assessing the complexity of localization algorithms is fundamental to evaluate their suitability to estimate the UE position with low-latency constraints. The proposed method for BI relies on a model obtained through adaptive boosting, whose complexity depends on the type of weak learners employed. In particular, the proposed method for BI leverages one-level CARTs, which determine a time complexity for training of $O(d\, N_{\mathrm{D}}\, N_{\mathrm{L}})$ and a runtime complexity of $O(d\, N_{\mathrm{L}})$, where $d = 10$ considering the vector of features described in Section III-A [69]. The low runtime complexity of the proposed method for BI makes it particularly suitable also for applications involving reduced capabilities devices, such as IoT and IIoT applications [74].

### E. BI Feature Selection

Feature selection is fundamental to optimize BI and discard less relevant features. For the proposed implementation of BI, which exploits one-level CARTs as weak learners, a feature relevance measure can be obtained by computing the relative frequency of the use of a feature as a splitting variable for the $N_{\mathrm{L}}$ weak learners [71]. Formally, the relevance of the $q$-th feature is given by

$$R_q = \frac{1}{N_{\mathrm{L}}} \sum_{m=1}^{N_{\mathrm{L}}} \mathbb{1}_{\{q\}}(\hat{q}^{[m]}) \tag{25}$$

where $\hat{q}^{[m]}$ is the index of the splitting variable exploited by the $m$-th one-level CART.




(a) InF-DH Case I (b) InF-DH Case II

Fig. 4. Examples of LOS maps for 3GPP (a) InF-DH Case I scenario; and (b) InF-DH Case II scenario. The colormap indicates the number of gNBs in LOS conditions for each position in the map and the red annuluses indicate the gNBs positions. The coordinates on the axis are in meters.

### *F. Integration of BI in SVE-Based Localization*

BI can be easily integrated into SVE-based localization algorithms. For example, the LS algorithm described in Section II-B can be transformed into a weighted least squares (WLS) algorithm modifying (2) as follows

$$\hat{\boldsymbol{p}} = \underset{\tilde{\boldsymbol{p}}}{\arg\min}\ \big(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\tilde{\boldsymbol{p}})\big)^{\mathrm{T}} \big(\boldsymbol{I} - \boldsymbol{\Psi}\big) \big(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}(\tilde{\boldsymbol{p}})\big) \tag{26}$$

where

$$\boldsymbol{\Psi} = \begin{bmatrix} \psi(\boldsymbol{\nu}_1) & 0 & \dots & 0 \\ 0 & \psi(\boldsymbol{\nu}_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \psi(\boldsymbol{\nu}_{N_\mathrm{g}}) \end{bmatrix} \tag{27}$$

and $\boldsymbol{I}$ is a $N_\mathrm{g} \times N_\mathrm{g}$ identity matrix [75], [76]. The use of $\boldsymbol{\Psi}$ enables to weight the contribution of each SVE in $\{\hat{\theta}_j\}_{j \in \mathcal{N}_\mathrm{g}}$ using a value equal to the estimated probability that it was obtained in LOS conditions.

### *G. Integration of BI in SI-Based Localization*

The information provided by NLOS identification can be integrated into SI-based localization. For example, a binary NLOS indicator, such as the one provided by DBI, can be integrated into SI-based localization determining two different generative models for NLOS and LOS conditions [43]. Specifically, let $\hat{\psi}$ be the DBI, with $\hat{\psi} = +1$ and $\hat{\psi} = -1$ respectively representing the identification of NLOS and LOS conditions. Also, let the error probability in detecting propagation conditions be defined by

$$\epsilon_{+1} = \mathbb{P}\big\{\mathrm{NLOS} \mid \hat{\psi} = -1\big\} \tag{28a}$$

$$\epsilon_{-1} = \mathbb{P}\big\{\mathrm{LOS} \mid \hat{\psi} = +1\big\}\,. \tag{28b}$$

The corresponding SFI is then given by

$$\mathcal{L}_{\hat{\theta}}(\theta) \propto \begin{cases} \epsilon_{-1}\,\mathcal{L}_\mathrm{L}(\theta) + (1-\epsilon_{-1})\,\mathcal{L}_\mathrm{NL}(\theta), & \text{for } \hat{\psi} = +1 \\ (1-\epsilon_{+1})\,\mathcal{L}_\mathrm{L}(\theta) + \epsilon_{+1}\,\mathcal{L}_\mathrm{NL}(\theta), & \text{for } \hat{\psi} = -1 \end{cases} \tag{29}$$

where $\mathcal{L}_\mathrm{NL}(\theta)$ and $\mathcal{L}_\mathrm{L}(\theta)$ are SFIs obtained from generative models tailored to NLOS and LOS propagation conditions, respectively. However, the approach based on (29) has several limitations. First of all, conventional NLOS identification categorizes the possible propagation conditions as binary, without taking into consideration the different wireless channel characteristics generated by NLOS propagation. Secondly, since the models are tailored to one propagation condition, errors in NLOS identification may harm localization accuracy.

The use of BI can effectively address such issues. In particular, BI can be integrated directly in the SFI as

$$\mathcal{L}_{\hat{\theta},\psi(\boldsymbol{\nu})}(\theta) \propto f_{\hat{\theta},\psi(\boldsymbol{\nu})}(\hat{\theta}, \psi(\boldsymbol{\nu}); \theta)\,. \tag{30}$$

Such approach allows modeling the SFI considering also the information provided by BI. Thus, the single generative model can describe a wider variety of situations and also mitigate errors that may occur in the information provided by the BI via the learning process. Moreover, the use of a single model for localization enables reducing the time needed for the UE localization with respect to the approach in (29). Note that the use of BI can also be interpreted as a two-step dimensionality reduction in SI-based localization. Specifically, the high-dimensional received RS is first processed to obtain a low-dimensional vector of features and then exploited to obtain BI for computing a one-dimensional probabilistic indicator that encapsulates rich positional information. Accordingly, the information encapsulated by BI can be intended not only as a probabilistic characterization of NLOS propagation conditions, but also as a normalized indication of the wireless channel quality.

## IV. Case Studies

This section quantifies the performance gain enabled by BI to both SVE-based localization and SI-based localization in complex wireless environments. We evaluate the localization performance in an industrial factory environment, namely the 3GPP-standardized InF-dense high (DH) scenario. The wireless propagation conditions of such scenario are modeled to represent a cluttered environment, characterized by a high density of metallic machinery with irregular structures [77]. Therefore, the InF-DH is a complex wireless scenario, with critical NLOS conditions and multipath propagation.

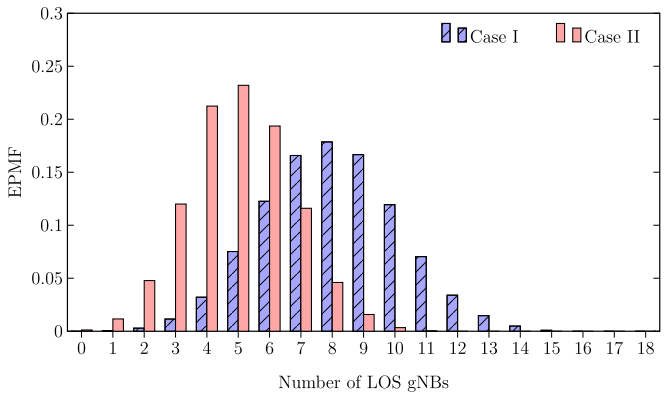



Fig. 5. EPMF of the number of wireless links in LOS conditions for the two cases considering a UE randomly deployed in the scenario.

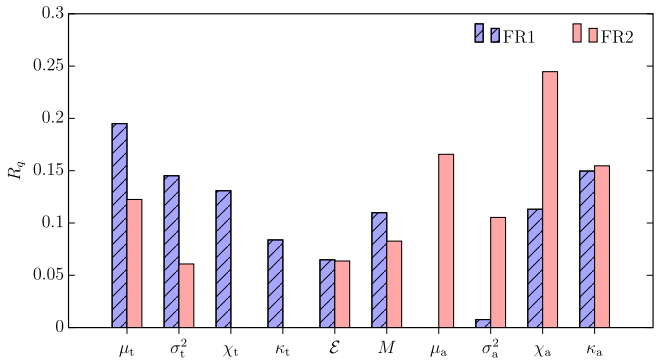



Fig. 6. Histogram of the feature relevance for the proposed BI method in the 3GPP-standardized InF-DH scenario for RSs transmitted in FR1 and FR2.

TABLE I
MAIN SIMULATION PARAMETERS

| **Paramter** | **FR1** | **FR2** |
|---|---|---|
| Carrier Frequency | 3.5 GHz | 28 GHz |
| Subcarrier Spacing | 30 kHz | 120 kHz |
| RSs Bandwidth | 100 MHz | 400 MHz |
| gNB Antenna Height | 8 m | 8 m |
| gNB Transmitted Power | 24 dBm | 24 dBm |
| gNB EIRP Limit | – | 58 dBm |
| gNB Noise Figure | 5 dB | 7 dB |
| gNB Antenna Radiation Pattern | [39] | [39] |
| UE Antenna Height | 1.5 m | 1.5 m |
| UE Transmitted Power | 23 dBm | 23 dBm |
| UE EIRP Limit | – | 43 dBm |
| UE Noise Figure | 9 dB | 13 dB |
| UE Antenna Radiation Pattern | [39] | [39] |

Results are reported in terms of empirical cumulative distribution function (ECDF) $\breve{F}(e_{\rm h})$ of the horizontal localization error $e_{\rm h}$, for two different case studies in the InF-DH scenario. Specifically, Case I consists of the baseline configuration in [39] with 18 gNBs, while Case II considers a different deployment where only 12 gNBs of the original layout are available. Fig. 4a and Fig. 4b show the gNBs deployment for the two case studies considered. The background of Fig. 4 illustrates an example of LOS map (i.e., a map which represents the number of gNBs in LOS conditions for each position), in the two case studies considered. Moreover, Fig. 5 reports the empirical probability mass function (EPMF) of the number of gNBs in LOS for a UE randomly deployed in the two cases. The localization performance is evaluated in FR1 and FR2 with carrier frequencies of 3.5 GHz and 28 GHz, respectively [39]. The bandwidths of the PRS and of the SRS are 100 MHz in FR1 and 400 MHz in FR2. The SVEs considered for localization in both FR1 and FR2 are DL-TOA, UL-TOA and RTT. AOD measurements are also considered in FR2 [39].

Results are obtained in conditions fully compliant with 3GPP specifications. This enables to evaluate the performance gain provided by BI under common 3GPP settings considered by companies and research institutions for comparing localization in 5G and beyond wireless networks. Specifically, the RSs are generated according to [51] and [78], and all the RSs parameters, including time-frequency allocation, periodicity, and sequence allocation are implemented according to [39]. The InF-DH wireless channels are generated based on the statistical models defined in [77] through the QuaDRiGa channel simulator [79], [80]. The NLOS conditions in the scenario are determined according to the NLOS probability model for the InF-DH scenario [77]. Moreover, both the wireless channel and the NLOS conditions are generated with spatial consistency according to [77]. All the setting parameters, including, transmitted power, equivalent isotropically radiated power (EIRP) and radiation patterns of both UEs and gNBs are set according to [39] and summarized in Table I. For each scenario setting, 300 instantiations of the scenario are generated, and for each of them, 10 UEs are deployed in the environment with random positions and orientations. The BI is obtained according to Section III-B using the set of features proposed in Section III-A. Specifically, the Real AdaBoost algorithm is implemented considering $N_{\rm L} = 600$ one-level CARTs as weak learners fit according to Section III-C.

Fig. 6 shows the feature relevance for the different features exploited for BI in the InF-DH scenario according to the method described in Section III-E. It can be observed that the features contributing to BI are different in FR1 and FR2. Specifically, time-based features are more effective in FR1 while amplitude-based features are more effective in FR2. Note that such analysis is scenario-dependent and therefore different environments may determine different feature relevances.

The localization performance is evaluated by making use of a 10-fold cross-validation technique [66]. Specifically, the data used in the training phase are split as follows:

- 70% of the training data are used for training the BI model, and identifying the number of iterations $N_{\rm I}$ that minimizes the ranging error in TOA estimation according to Section II-A. While the selection of the best $N_{\rm I}$ is highly beneficial for SVE-based localization, in SI-based localization such measurement refinement can be irrelevant or even detrimental given the ML-based approach

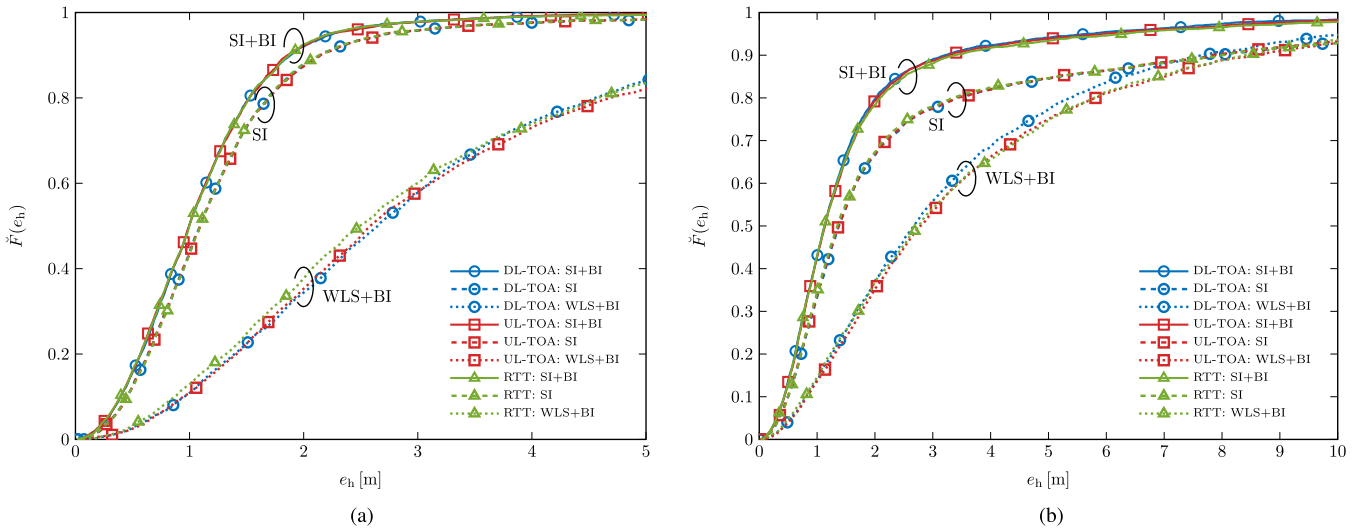



Fig. 7. Localization performance in FR1. The performance is reported in terms of ECDF of the horizontal localization error for SI-based localization with BI, SI-based localization without BI, and WLS-based localization with BI. The results are reported for (a) Case I; and (b) Case II.

of the algorithm. Therefore, for SI-based localization, a simpler TOA estimation with $N_{\mathrm{I}} = 1$ is considered.

- 30% of the training data are used for determining the generative models for SI-based localization, which consist of GMMs with 12 components, and estimating the DBI error probabilities $\epsilon_{+1}$ and $\epsilon_{-1}$ for SI-based localization with conventional NLOS identification according to (29).

The data not exploited in the training phase are used in the online phase to quantify the performance gain provided by BI to SI-based localization and SVE-based localization considering the different measurements described in Section II-A. Results for localization based on the fusion of DL-TOA and AOD, hereafter referred to as DL-FUS, are also reported.

### A. Results in FR1

Fig. 7a shows the ECDF of the horizontal localization error in FR1 for the Case I of the InF-DH scenario. It can be observed that WLS-based localization with BI provides an accuracy of around 6 m at the 90th percentile for all the measurements considered. Moreover, it can be observed that conventional SI-based localization outperforms WLS-based localization for all the configurations considered. The use of BI in SI-based localization enables a further improvement. In particular, at the 99th percentile, the localization error is reduced from around 7 m to less than 4 m for all the measurements considered. Note that very similar performance is obtained with localization based on DL-TOA and on UL-TOA despite the lowest transmitted power of SRS. Similar performance is also obtained considering localization based on RTT.

Fig. 7b shows the ECDF of the horizontal localization error in FR1 for the Case II of the InF-DH scenario. It can be observed that WLS-based localization with BI provides performance comparable to the ones obtained with SI-based localization at the percentiles over the 90th for all the measurements considered. However, if BI is integrated into SI, the localization error is reduced from around 8 m to less than 3.5 m at the 90th percentile.

### B. Results in FR2

Fig. 8a and Fig. 8b show the ECDFs of the horizontal localization error in FR2 for the Case I of the InF-DH scenario. It can be observed that SI-based localization with BI considering DL-TOA, UL-TOA, and RTT measurements provide similar perfrormance. Specifically, they achieve around 1 m and 1.5 m localization error at the 80th and at the 90th percentile, respectively. Moreover, it can be noticed that also in FR2 SI-based localization with BI outperforms WLS-based localization with BI, providing an accuracy improvement of approximately 8 m for both DL-TOA and UL-TOA at the 90th percentile. Such performance improvement is even more noticeable at the 99th percentile, where the use of BI in SI-based localization provides a performance gain of approximately 10 m for both localization using DL-TOA and UL-TOA with respect to WLS-based localization with BI. Similar performance is obtained also for localization based on RTT measurements. It can also be noticed that localization based on AOD measurements determines approximately 3 m of horizontal localization error at the 80th percentile using SI-based localization with BI, which represents a significant improvement when compared to the performance provided by WLS-based localization with BI. Considering SI-based localization with BI and fusing DL-TOA and AOD measurements, a significant localization accuracy improvement at every percentile is obtained with respect localization using the single measurements. In particular, such approach enables sub-meter localization accuracy at the 80th percentile and 1.56 m and 2.67 m of horizontal localization error at the 95th and 99th percentile, respectively. Lastly, it can be observed that SI-based localization with BI provides a significant performance gain with respect to SI-based localization without BI (Table III). For example, considering localization with DL-TOA measurements, it can be observed that BI enables

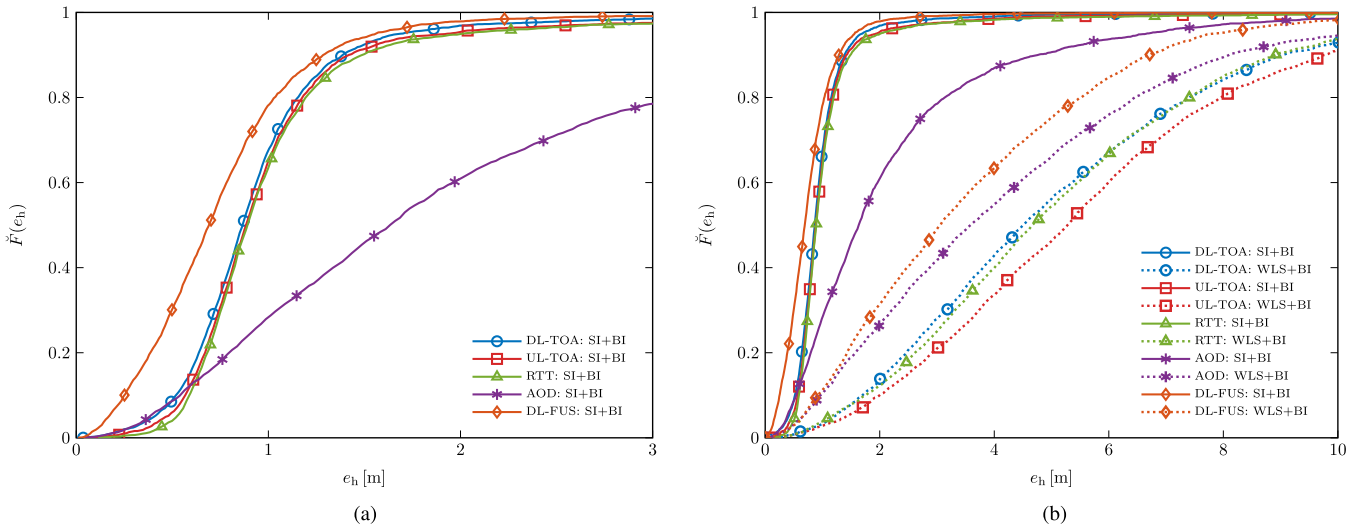



Fig. 8. Localization performance for Case I in FR2. The performance is reported in terms of ECDF of the horizontal localization error for (a) SI-based localization with BI; and (b) SI-based localization with BI versus WLS-based localization with BI.

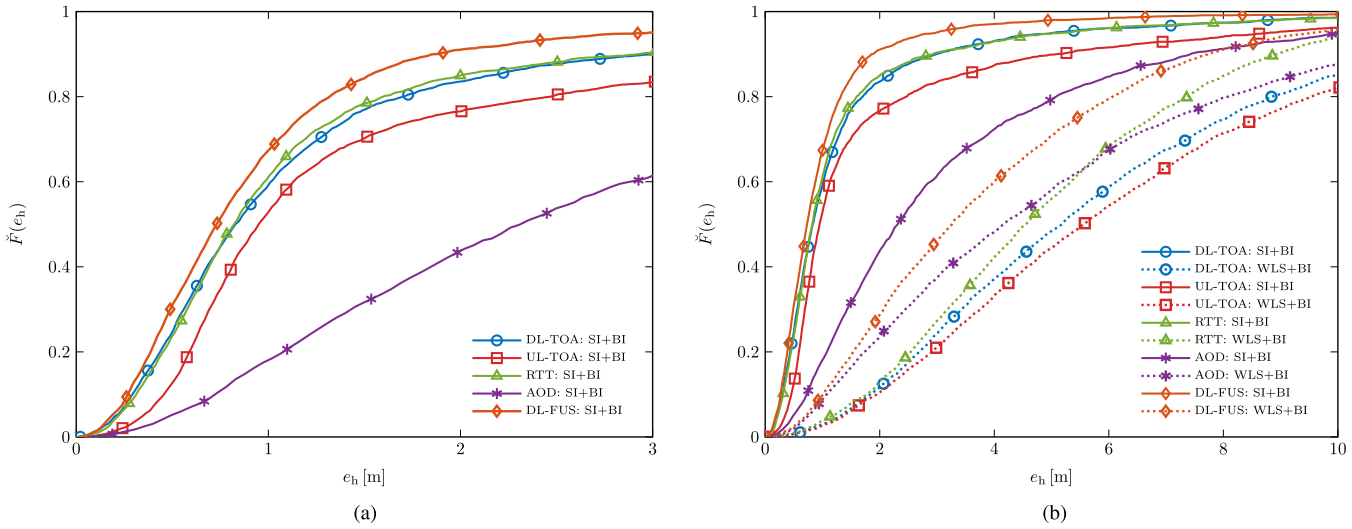



Fig. 9. Localization performance for Case II in FR2. The performance is reported in terms of ECDF of the horizontal localization error for (a) SI-based localization with BI; and (b) SI-based localization with BI versus WLS-based localization with BI.

a performance gain of approximatively 2 m and 11 m at the 95th and at the 99th percentile, respectively.

Fig. 9a and Fig. 9b show the ECDFs of the horizontal localization error in FR2 for the Case II of the InF-DH scenario. It can be observed that SI-based localization with BI considering DL-TOA and RTT measurements provide similar perfrormance. Specifically, they achieve around 1.5 m and 3 m localization error at the 80th and at the 90th percentile, respectively. Moreover, it can be noticed that at the 95th percentile, the use of BI in SI-based localization provides a performance gain of approximatively 7 m for localization using DL-TOA or RTT measurements with respect to WLS-based localization with BI. However, if SI-based localization with BI is obtained considering UL-TOA measurements, the localization accuracy is around 2.5 m and 5 m at the 80th and at the 90th percentile, respectively. It can also be noticed that localization based on AOD measurements determines approximately 5 m of horizontal localization error at the 80th percentile using SI-based localization with BI. This represents a significant improvement when compared to the 9 m achieved by WLS-based localization with BI at the same percentile. Similarly to Case I, it can be observed that the fusion of DL-TOA and AOD measurements is beneficial to the localization performance. In particular, with this configuration the localization error is around 1.2 m at the 80th percentile and equal to 1.87 m and 2.98 m at the 90th and 95th percentile, respectively. Lastly, it can be observed that SI-based localization with BI provides a significant performance gain with respect to SI-based localization without BI (Table III). For example, considering localization with RTT measurements, it can be observed that BI enables a performance gain of approximatively 7 m at both the 95th and at the 99th percentile.

For all the configurations considered, it can be observed that localization in Case II provides worse performance when compared to Case I. This is due to the reduced number of gNBs and becomes particularly relevant with UL localization

TABLE II
RELEVANT LOCALIZATION ERROR PERCENTILES IN FR1 AND PERCENTUAL ACCURACY GAIN WITH RESPECT TO SI-BASED LOCALIZATION (–)

| | Case I | | | | | | Case II | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Configuration** | 90th | | 95th | | 99th | | 90th | | 95th | | 99th | |
| DL-TOA: LS | 14.31 m | (-572%) | 16.86 m | (-591%) | 24.58 m | (-249%) | 16.07 m | (-103%) | 19.28 m | (-68%) | 28.34 m | (-53%) |
| DL-TOA: WLS+BI | 6.01 m | (-182%) | 7.31 m | (-168%) | 10.43 m | (-48%) | 7.64 m | (+3%) | 10.12 m | (+12%) | 16.11 m | (+13%) |
| DL-TOA: SI | 2.13 m | – | 2.73 m | – | 7.04 m | – | 7.90 m | – | 11.50 m | – | 18.57 m | – |
| DL-TOA: SI+DBI | 2.07 m | (+3%) | 2.78 m | (-2%) | 9.42 m | (-34%) | 6.38 m | (+19%) | 10.32 m | (+10%) | 17.09 m | (+8%) |
| DL-TOA: SI+BI | 1.89 m | (+11%) | 2.25 m | (+18%) | 3.88 m | (+45%) | 3.20 m | (+59%) | 5.67 m | (+51%) | 11.22 m | (+40%) |
| UL-TOA: LS | 14.97 m | (-596%) | 17.43 m | (-543%) | 25.01 m | (-254%) | 16.61 m | (-112%) | 20.18 m | (-75%) | 28.99 m | (-54%) |
| UL-TOA: WLS+BI | 6.47 m | (-201%) | 8.96 m | (-231%) | 23.59 m | (-234%) | 8.59 m | (-10%) | 11.44 m | (+1%) | 26.60 m | (-41%) |
| UL-TOA: SI | 2.15 m | – | 2.71 m | – | 7.07 m | - | 7.84 m | – | 11.50 m | – | 18.84 m | – |
| UL-TOA: SI+DBI | 2.09 m | (+3%) | 2.80 m | (-3%) | 9.48 m | (-34%) | 6.40 m | (+19%) | 10.34 m | (+10%) | 16.91 m | (+10%) |
| UL-TOA: SI+BI | 1.85 m | (+14%) | 2.21 m | (+18%) | 3.79 m | (+46%) | 3.29 m | (+58%) | 5.93 m | (+48%) | 12.43 m | (+34%) |
| RTT: LS | 14.57 m | (-568%) | 17.07 m | (-532%) | 24.83 m | (-271%) | 16.31 m | (-103%) | 19.76 m | (-71%) | 28.65 m | (-52%) |
| RTT: WLS+BI | 6.13 m | (-189%) | 8.08 m | (-199%) | 12.38 m | (-85%) | 8.42 m | (-5%) | 11.08 m | (+4%) | 19.97 m | (-6%) |
| RTT: SI | 2.18 m | – | 2.70 m | – | 6.70 m | – | 8.03 m | – | 11.57 m | – | 18.91 m | – |
| RTT: SI+DBI | 2.08 m | (+5%) | 2.88 m | (-7%) | 10.19 m | (-52%) | 6.74 m | (+16%) | 10.78 m | (+7%) | 17.06 m | (+10%) |
| RTT: SI+BI | 1.87 m | (+14%) | 2.22 m | (+18%) | 3.84 m | (+43%) | 3.39 m | (+58%) | 6.32 m | (+45%) | 13.06 m | (+31%) |

due to the lower transmitted power of SRS. It can also be observed that in Case II BI is even more impactful than in Case I. This is because the total positional information provided by the SVEs is less with respect to Case I due to the reduced number of gNBs.

*Remark:* BI is more effective in FR2 than in FR1. This is because at high frequencies the signal attenuation due to NLOS conditions is higher, thus the positional information encapsulated by the SVEs is lower. Accordingly, BI plays a fundamental role in performing accurate localization, especially when the positional information provided by the SVEs is limited (e.g., in FR2 or with a reduced number of gNBs available for localization).

### C. BI Localization Performance Gain

This section evaluates the performance gain offered by BI. In particular, results are obtained for all the localization techniques considered in the previous sections, including LS-based localization according to (2), and conventional SI-based localization with DBI according to (29). The results are reported at the 90th, 95th, and 99th percentile both in terms of horizontal localization error and in terms of percentual gain with respect to conventional SI-based localization without BI.

Table II shows the localization performance in FR1. It can be observed that the use of BI together with SI-based localization enables a gain between 11% and 46% for Case I and between 31% and 59% for Case II when compared to conventional SI-based localization considering all the percentiles and measurements reported. In particular, the average localization accuracy gain is 25% for Case I and 47% for Case II. Moreover, it can be noticed that SI-based localization with DBI is not able to provide a performance gain, despite the over 90% NLOS identification accuracy of DBI, as reported in Fig. 10a.

Table III shows the localization performance in FR2. It can be observed that the use of BI together with SI-based localization enables a gain between 23% and 70% for Case I and between 18% and 68% for Case II when compared to conventional SI-based localization considering all the percentiles and measurements reported. In particular, the average localization accuracy gain is 43% for Case I and 40% in Case II. Moreover, it can be noticed that SI-based localization with DBI in almost every case provides worse performance with respect to conventional SI-based localization, despite the around 85% NLOS identification accuracy of DBI, as reported in Fig. 10b. This is due to the limitations described in Section III-G of the approach based on (29) and further demonstrates the importance of the probabilistic information provided by BI.

Lastly, it can be observed that for all the measurements considered, LS-based localization is not able to provide satisfactory localization performance in both FR1 and FR2.

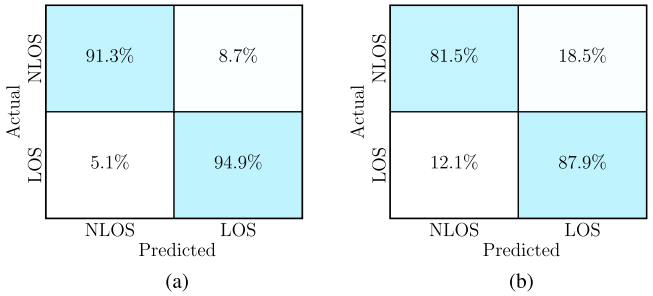



Fig. 10. Confusion matrix of DBI in (a) FR1; and (b) FR2.

## V. CONCLUSION

This paper introduced the concept of blockage intelligence (BI) to provide a probabilistic characterization of wireless propagation conditions. Then, it discusses its integration in both single-value estimate (SVE)-based and soft information (SI)-based localization algorithms. In particular, a low-complexity BI method for providing an indicator that encapsulates rich positional information on the non-line-of-sight (NLOS) conditions is proposed for fifth generation (5G) and beyond wireless networks. The performance gain provided by BI is evaluated for different localization algorithms in 3rd Generation Partnership Project (3GPP) industrial scenarios

TABLE III
RELEVANT LOCALIZATION ERROR PERCENTILES IN FR2 AND PERCENTUAL ACCURACY GAIN WITH RESPECT TO SI-BASED LOCALIZATION (–)

| | Case I | | | | | | Case II | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Configuration** | 90th | | 95th | | 99th | | 90th | | 95th | | 99th | |
| DL-TOA: LS | 15.71 m | (-740%) | 19.52 m | (-445%) | 29.43 m | (-129%) | 17.27 m | (-94%) | 21.83 m | (-74%) | 32.99 m | (-80%) |
| DL-TOA: WLS+BI | 9.04 m | (-383%) | 10.94 m | (-206%) | 14.83 m | (-16%) | 11.49 m | (-29%) | 14.17 m | (-13%) | 19.09 m | (-4%) |
| DL-TOA: SI | 1.87 m | – | 3.58 m | – | 12.87 m | – | 8.89 m | – | 12.57 m | – | 18.28 m | – |
| DL-TOA: SI+DBI | 3.63 m | (-94%) | 8.04 m | (-125%) | 14.72 m | (-14%) | 11.04 m | (-24%) | 14.26 m | (-13%) | 18.62 m | (-2%) |
| DL-TOA: SI+BI | 1.38 m | (+26%) | 1.68 m | (+53%) | 3.89 m | (+70%) | 3.00 m | (+66%) | 5.03 m | (+60%) | 11.01 m | (+40%) |
| UL-TOA: LS | 15.82 m | (-556%) | 18.59 m | (-304%) | 26.00 m | (-97%) | 17.07 m | (-77%) | 20.14 m | (-52%) | 28.20 m | (-53%) |
| UL-TOA: WLS+BI | 9.78 m | (-306%) | 11.32 m | (-146%) | 15.00 m | (-14%) | 12.25 m | (-27%) | 14.63 m | (-10%) | 21.00 m | (-14%) |
| UL-TOA: SI | 2.41 m | – | 4.60 m | – | 13.21 m | – | 9.62 m | – | 13.25 m | – | 18.49 m | – |
| UL-TOA: SI+DBI | 4.35 m | (-80%) | 8.70 m | (-89%) | 15.19 m | (-15%) | 11.21 m | (-17%) | 14.26 m | (-8%) | 18.88 m | (-2%) |
| UL-TOA: SI+BI | 1.41 m | (+41%) | 1.93 m | (+58%) | 4.66 m | (+62%) | 5.09 m | (+47%) | 8.78 m | (+34%) | 15.00 m | (+19%) |
| RTT: LS | 15.38 m | (-554%) | 17.99 m | (-309%) | 24.82 m | (-91%) | 16.56 m | (-82%) | 19.93 m | (-58%) | 27.16 m | (-48%) |
| RTT: WLS+BI | 8.91 m | (-279%) | 10.58 m | (-140%) | 14.29 m | (-10%) | 11.54 m | (-27%) | 13.83 m | (-10%) | 19.38 m | (-6%) |
| RTT: SI | 2.35 m | – | 4.40 m | – | 12.98 m | – | 9.12 m | – | 12.63 m | – | 18.33 m | – |
| RTT: SI+DBI | 3.77 m | (-60%) | 8.15 m | (-85%) | 15.73 m | (-21%) | 10.98 m | (-20%) | 14.05 m | (-11%) | 18.72 m | (-2%) |
| RTT: SI+BI | 1.48 m | (+37%) | 2.02 m | (+54%) | 6.23 m | (+52%) | 2.92 m | (+68%) | 5.21 m | (+59%) | 11.20 m | (+39%) |
| AOD: LS | 9.26 m | (-38%) | 11.12 m | (-20%) | 15.49 m | (-7%) | 12.99 m | (-19%) | 15.65 m | (-3%) | 22.46 m | (+2%) |
| AOD: WLS+BI | 8.10 m | (-21%) | 10.36 m | (-12%) | 15.24 m | (-6%) | 10.89 m | (+0%) | 14.30 m | (+6%) | 22.17 m | (+3%) |
| AOD: SI | 6.70 m | – | 9.25 m | – | 14.43 m | – | 10.91 m | – | 15.18 m | – | 22.88 m | – |
| AOD: SI+DBI | 5.51 m | (+18%) | 7.50 m | (+19%) | 12.45 m | (+14%) | 8.39 m | (+23%) | 11.42 m | (+25%) | 19.63 m | (+14%) |
| AOD: SI+BI | 4.75 m | (+29%) | 6.64 m | (+28%) | 11.16 m | (+23%) | 7.41 m | (+32%) | 10.22 m | (+33%) | 18.71 m | (+18%) |
| DL-FUS: LS | 8.77 m | (-407%) | 10.51 m | (-311%) | 13.81 m | (-186%) | 9.95 m | (-292%) | 12.05 m | (-157%) | 18.24 m | (-92%) |
| DL-FUS: WLS+BI | 6.69 m | (-287%) | 7.83 m | (-206%) | 10.82 m | (-124%) | 7.74 m | (-205%) | 9.56 m | (-104%) | 13.41 m | (-41%) |
| DL-FUS: SI | 1.73 m | – | 2.56 m | – | 4.83 m | – | 2.54 m | – | 4.69 m | – | 9.49 m | – |
| DL-FUS: SI+DBI | 1.93 m | (-12%) | 3.27 m | (-28%) | 8.78 m | (-82%) | 6.40 m | (-152%) | 8.94 m | (-91%) | 14.09 m | (-48%) |
| DL-FUS: SI+BI | 1.28 m | (+26%) | 1.56 m | (+39%) | 2.67 m | (+45%) | 1.87 m | (+26%) | 2.98 m | (+36%) | 7.35 m | (+23%) |

and compared with conventional binary NLOS identification. Results show that BI provides significant performance gain to both SVE-based and SI-based localization. In particular, BI together with SI-based localization is able to largely outperform conventional localization approaches in both frequency range 1 (FR1) and frequency range 2 (FR2). Indeed, BI provides rich positional information which can be leveraged to enhance location awareness in complex wireless environments. Results in 3GPP indoor factory scenarios show that BI is a key enabler to achieve accurate localization in 5G and beyond wireless networks operating in complex wireless environments.

## ACKNOWLEDGMENT

The authors wish to thank C. A. Gómez-Vega and F. Morselli for their helpful suggestions and careful reading of the manuscript.

## REFERENCES

[1] M. Win et al., "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.

[2] W. Chen, J. Montojo, J. Lee, M. Shafi, and Y. Kim, "The standardization of 5G-advanced in 3GPP," *IEEE Commun. Mag.*, vol. 60, no. 11, pp. 98–104, Nov. 2022.

[3] M. Z. Win, W. Dai, Y. Shen, G. Chrisikos, and H. V. Poor, "Network operation strategies for efficient localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1224–1254, Jul. 2018.

[4] *Technical Specification Group Radio Access Network; Technical Specification Group Services and System Aspects; Study on Positioning use Cases*, document TR 22.872 V16.1.0, Release 16, 3GPP, Sep. 2018.

[5] A. Conti et al., "Location awareness in beyond 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 22–27, Nov. 2021.

[6] J. Thomas, J. Welde, G. Loianno, K. Daniilidis, and V. Kumar, "Autonomous flight for detection, localization, and tracking of moving targets with a small quadrotor," *IEEE Robot. Autom. Lett.*, vol. 2, no. 3, pp. 1762–1769, Jul. 2017.

[7] R. Karlsson and F. Gustafsson, "The future of automotive localization algorithms: Available, reliable, and scalable localization: Anywhere and anytime," *IEEE Signal Process. Mag.*, vol. 34, no. 2, pp. 60–69, Mar. 2017.

[8] G. Zhan and W. Shi, "LOBOT: Low-cost, self-contained localization of small-sized ground robotic vehicles," *IEEE Trans. Parallel Distrib. Syst.*, vol. 24, no. 4, pp. 744–753, Apr. 2013.

[9] D. Wu, D. Chatzigeorgiou, K. Youcef-Toumi, and R. Ben-Mansour, "Node localization in robotic sensor networks for pipeline inspection," *IEEE Trans. Ind. Informat.*, vol. 12, no. 2, pp. 809–819, Apr. 2016.

[10] M. Mozaffari, W. Saad, M. Bennis, Y. Nam, and M. Debbah, "A tutorial on UAVs for wireless networks: Applications, challenges, and open problems," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2334–2360, 3rd Quart., 2019.

[11] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[12] V. Moreno, M. A. Zamora, and A. F. Skarmeta, "A low-cost indoor localization system for energy sustainability in smart buildings," *IEEE Sensors J.*, vol. 16, no. 9, pp. 3246–3262, May 2016.

[13] K. Lin, M. Chen, J. Deng, M. M. Hassan, and G. Fortino, "Enhanced fingerprinting and trajectory prediction for IoT localization in smart buildings," *IEEE Trans. Autom. Sci. Eng.*, vol. 13, no. 3, pp. 1294–1307, Jul. 2016.

[14] F. Zabini and A. Conti, "Inhomogeneous Poisson sampling of finite-energy signals with uncertainties in $\mathbb{R}^d$," *IEEE Trans. Signal Process.*, vol. 64, no. 18, pp. 4679–4694, Sep. 2016.

[15] M. Di Renzo et al., "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[16] J. Ji, A. Khajepour, W. W. Melek, and Y. Huang, "Path planning and tracking for vehicle collision avoidance based on model predictive control with multiconstraints," *IEEE Trans. Veh. Technol.*, vol. 66, no. 2, pp. 952–964, Feb. 2017.

[17] M. Driusso, C. Marshall, M. Sabathy, F. Knutti, H. Mathis, and F. Babich, "Vehicular position tracking using LTE signals," *IEEE Trans. Veh. Technol.*, vol. 66, no. 4, pp. 3376–3391, Apr. 2017.

[18] K. Witrisal et al., "High-accuracy localization for assisted living: 5G systems will turn multipath channels from foe to friend," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 59–70, Mar. 2016.

[19] M. Chiani, A. Giorgetti, and E. Paolini, "Sensor radar for object tracking," *Proc. IEEE*, vol. 106, no. 6, pp. 1022–1041, Jun. 2018.

[20] H. Wymeersch, G. Seco-Granados, G. Destino, D. Dardari, and F. Tufvesson, "5G mmWave positioning for vehicular networks," *IEEE Wireless Commun.*, vol. 24, no. 6, pp. 80–86, Dec. 2017.

[21] S. D'Oro, L. Galluccio, G. Morabito, and S. Palazzo, "Exploiting object group localization in the Internet of Things: Performance analysis," *IEEE Trans. Veh. Technol.*, vol. 64, no. 8, pp. 3645–3656, Aug. 2015.

[22] S. G. Nagarajan, P. Zhang, and I. Nevat, "Geo-spatial location estimation for Internet of Things (IoT) networks with one-way time-of-arrival via stochastic censoring," *IEEE Internet Things J.*, vol. 4, no. 1, pp. 205–214, Feb. 2017.

[23] A. A. Abdellatif, A. Mohamed, C. F. Chiasserini, M. Tlili, and A. Erbad, "Edge computing for smart health: Context-aware approaches, opportunities, and challenges," *IEEE Netw.*, vol. 33, no. 3, pp. 196–203, May 2019.

[24] L. Catarinucci et al., "An IoT-aware architecture for smart healthcare systems," *IEEE Internet Things J.*, vol. 2, no. 6, pp. 515–526, Dec. 2015.

[25] E. S. Lohan et al., "Benefits of positioning-aided communication technology in high-frequency industrial IoT," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 142–148, Dec. 2018.

[26] S. Vitturi, C. Zunino, and T. Sauter, "Industrial communication systems and their future challenges: Next-generation Ethernet, IIoT, and 5G," *Proc. IEEE*, vol. 107, no. 6, pp. 944–961, Jun. 2019.

[27] M. Posluk, J. Ahlander, D. Shrestha, S. M. Razavi, G. Lindmark, and F. Gunnarsson, "5G deployment strategies for high positioning accuracy in indoor environments," in *Proc. Int. Conf. Indoor Positioning Indoor Navigat. (IPIN)*, Lloret de Mar, Spain, 2021. [Online]. Available: https://ceur-ws.org/Vol-3097/

[28] M. Elsanhoury et al., "Precision positioning for smart logistics using ultra-wideband technology-based indoor navigation: A review," *IEEE Access*, vol. 10, pp. 44413–44445, 2022.

[29] *Technical Specification Group Services and System Aspects; Service Requirements for the 5G System; Stage 1*, document TS 22.261 V19.1.0, Release 19, 3GPP, Dec. 2022.

[30] T. Jiang et al., "3GPP standardized 5G channel model for IIoT scenarios: A survey," *IEEE Internet Things J.*, vol. 8, no. 11, pp. 8799–8815, Jun. 2021.

[31] J. Palacios, G. Bielsa, P. Casari, and J. Widmer, "Single- and multiple-access point indoor localization for millimeter-wave networks," *IEEE Trans. Wireless Commun.*, vol. 18, no. 3, pp. 1927–1942, Mar. 2019.

[32] S. Sun, T. S. Rappaport, M. Shafi, P. Tang, J. Zhang, and P. J. Smith, "Propagation models and performance evaluation for 5G millimeter-wave bands," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8422–8439, Sep. 2018.

[33] M. Shafi et al., "Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems," *IEEE Commun. Mag.*, vol. 56, no. 12, pp. 14–20, Dec. 2018.

[34] *Technical Specification Group Radio Access Network; NR; Physical Layer Measurements*, document TS 38.215 V17.2.0, 3GPP, Release 17, Sep. 2022.

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

[36] S. Dwivedi et al., "Positioning in 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 38–44, Nov. 2021.

[37] S. Marano, W. Gifford, H. Wymeersch, and M. Win, "NLOS identification and mitigation for localization based on UWB experimental data," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 7, pp. 1026–1035, Sep. 2010.

[38] A. Conti, M. Guerra, D. Dardari, N. Decarli, and M. Z. Win, "Network experimentation for cooperative localization," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 467–475, Feb. 2012.

[39] *Technical Specification Group Radio Access Network; Study on NR Positioning Enhancements*, document TR 38.857 V17.0.0, 3GPP, Release 17, Mar. 2021.

[40] *Technical Specification Group Radio Access Network; Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR Air Interface*, document TR 38.843 V18.0.0, 3GPP, Release 18, Jan. 2022.

[41] X. Lin, "An overview of 5G advanced evolution in 3GPP release 18," *IEEE Commun. Standards Mag.*, vol. 6, no. 3, pp. 77–83, Sep. 2022.

[42] X. Lin et al., "Fueling the next quantum leap in cellular networks: Embracing AI in 5G evolution towards 6G," Nov. 2021, *arXiv:2111.10663*.

[43] A. Conti, S. Mazuelas, S. Bartoletti, W. C. Lindsey, and M. Z. Win, "Soft information for Localization-of-Things," *Proc. IEEE*, vol. 107, no. 11, pp. 2240–2264, Nov. 2019.

[44] F. Morselli, S. M. Razavi, M. Z. Win, and A. Conti, "Soft information based localization for 5G networks and beyond," *IEEE Trans. Wireless Commun.*, early access, May 16, 2023.

[45] K. Yu, K. Wen, Y. Li, S. Zhang, and K. Zhang, "A novel NLOS mitigation algorithm for UWB localization in harsh indoor environments," *IEEE Trans. Veh. Technol.*, vol. 68, no. 1, pp. 686–699, Jan. 2019.

[46] B. Silva and G. P. Hancke, "IR-UWB-based non-line-of-sight identification in harsh environments: Principles and challenges," *IEEE Trans. Ind. Informat.*, vol. 12, no. 3, pp. 1188–1195, Jun. 2016.

[47] B. Silva and G. P. Hancke, "Ranging error mitigation for through-the-wall non-line-of-sight conditions," *IEEE Trans. Ind. Informat.*, vol. 16, no. 11, pp. 6903–6911, Nov. 2020.

[48] K. Yu and Y. J. Guo, "Statistical NLOS identification based on AOA, TOA, and signal strength," *IEEE Trans. Veh. Technol.*, vol. 58, no. 1, pp. 274–286, Jan. 2009.

[49] *Technical Specification Group Radio Access Network; NG-RAN; NR Positioning Protocol A (NRPPa)*, document TS 38.455 V17.3.0, Release 17, 3GPP, Jan. 2023.

[50] *Technical Specification Group Radio Access Network; NG Radio Access Network (NG-RAN); Stage 2 Functional specification of User Equipment (UE) Positioning in NG-RAN*, document TS 38.305 V17.3.0, Release 17, 3GPP, Jan. 2023s.

[51] *Technical Specification Group Radio Access Network; NR; Physical Channels and Modulation*, document TS 38.211 V17.4.0, 3GPP, Release 17, Jan. 2021.

[52] G. Torsoli, M. Z. Win, and A. Conti, "Selection of reference base station for TDOA-based localization in 5G and beyond IIoT," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Rio de Janeiro, Brazil, Dec. 2022, pp. 317–322.

[53] A. F. Molisch, "Ultra-wide-band propagation channels," *Proc. IEEE*, vol. 97, no. 2, pp. 353–371, Feb. 2009.

[54] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proc. IEEE*, vol. 97, no. 2, pp. 404–426, Feb. 2009.

[55] H. Ryden, A. A. Zaidi, S. M. Razavi, F. Gunnarsson, and I. Siomina, "Enhanced time of arrival estimation and quantization for positioning in LTE networks," in *Proc. IEEE 27th Annu. Int. Symp. Pers., Indoor, Mobile Radio Commun. (PIMRC)*, Valencia, Spain, Sep. 2016, pp. 1–6.

[56] *Technical Specification Group Services and System Aspects; System Architecture for the 5G System (5GS); Stage 2*, document TS 23.501 V18.0.0, 3GPP, Release 18, Dec. 2022.

[57] H. Holma, A. Toskala, and T. Nakamura, *5G Technology: 3GPP New Radio*. Hoboken, NJ, USA: Wiley, 2020.

[58] *Integration of 5G With Time-Sensitive Networking for Industrial Communications*, 5G-ACIA, Feb. 2021. [Online]. Available: https://www.5g-acia.org

[59] M. Wen et al., "Private 5G networks: Concepts, architectures, and research landscape," *IEEE J. Sel. Topics Signal Process.*, vol. 16, no. 1, pp. 7–25, Jan. 2022.

[60] A. Rostami, "Private 5G networks for vertical industries: Deployment and operation models," in *Proc. IEEE 2nd 5G World Forum (5GWF)*, Sep. 2019, pp. 433–439.

[61] Y.-N. R. Li, B. Gao, X. Zhang, and K. Huang, "Beam management in millimeter-wave communications for 5G and beyond," *IEEE Access*, vol. 8, pp. 13282–13293, 2020.

[62] C. Mensing and S. Plass, "Positioning algorithms for cellular networks using TDOA," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, vol. 4, May 2006, p. 4.

[63] H. Rydén et al., "Baseline performance of LTE positioning in 3GPP 3D MIMO indoor user scenarios," in *Proc. Int. Conf. Localization GNSS (ICL-GNSS)*, Gothenburg, Sweden, Jun. 2015, pp. 1–6.

[64] F. Gustafsson and F. Gunnarsson, "Mobile positioning using wireless networks: Possibilities and fundamental limitations based on available wireless network measurements," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 41–53, Jul. 2005.
[65] J. Klemelä, *Smoothing of Multivariate Data: Density Estimation and Visualization*. Hoboken, NJ, USA: Wiley, 2009.
[66] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
[67] S. Bartoletti, A. Giorgetti, M. Z. Win, and A. Conti, "Blind selection of representative observations for sensor radar networks," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1388–1400, Apr. 2015.
[68] I. Guvenc, C.-C. Chong, and F. Watanabe, "NLOS identification and mitigation for UWB localization systems," in *Proc. IEEE Wireless Commun. Netw. Conf.*, Hong Kong, Mar. 2007, pp. 1571–1576.
[69] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning*, 2nd ed. New York, NY, USA: Springer, 2009.
[70] J. Friedman, T. Hastie, and R. Tibshirani, "Additive logistic regression: A statistical view of boosting," *Ann. Statist.*, vol. 28, no. 2, pp. 337–407, Apr. 2000.
[71] L. Breiman, J. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Belmont, CA, USA: Wadsworth Publishing Company, 1984.
[72] R. E. Schapire, "The strength of weak learnability," *Mach. Learn.*, vol. 5, no. 2, pp. 197–227, Jun. 1990.
[73] T. Hastie, S. Rosset, J. Zhu, and H. Zou, "Multi-class AdaBoost," *Statist. Its Interface*, vol. 2, no. 3, pp. 349–360, 2009.
[74] M. Vaezi et al., "Cellular, wide-area, and non-terrestrial IoT: A survey on 5G advances and the road toward 6G," *IEEE Commun. Surveys Tuts.*, vol. 24, no. 2, pp. 1117–1174, 2nd Quart., 2022.
[75] İ. Güvenç, C.-C. Chong, F. Watanabe, and H. Inamura, "NLOS identification and weighted least-squares localization for UWB systems using multipath channel statistics," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, pp. 1–14, Dec. 2007.
[76] D. Dardari, A. Conti, J. Lien, and M. Z. Win, "The effect of cooperation on UWB-based positioning systems using experimental data," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, pp. 1–11, Dec. 2008.
[77] *Technical Specification Group Radio Access Network; Study on Channel Model for Frequencies From 0.5 to 100 GHz*, document TR 38.901 V17.0.0, 3GPP, Mar. 2022, Release 17.
[78] *Technical Specification Group Radio Access Network; NG-RAN; Physical Layer Procedures for Data*, document TS 38.214 V17.4.0, 3GPP, Release 17, Jan. 2023.
[79] S. Jaeckel, L. Raschkowski, K. Borner, and L. Thiele, "QuaDRiGa: A 3-D multi-cell channel model with time evolution for enabling virtual field trials," *IEEE Trans. Antennas Propag.*, vol. 62, no. 6, pp. 3242–3256, Jun. 2014.
[80] S. Jaeckel, L. Raschkowski, K. Börner, L. Thiele, F. Burkhardt, and E. Eberlein, "QuaDRiGa—Quasi deterministic radio channel generator, user manual and documentation," Fraunhofer Heinrich Hertz Inst., Tech. Rep. v2.6.1, 2021.

**Gianluca Torsoli** (Student Member, IEEE) received the Laurea degree (*summa cum laude*) in Electronics for ICT engineering in 2021 from the University of Ferrara, Italy. He is currently pursuing the Ph.D. degree with the Wireless Communication and Localization Networks Laboratory, University of Ferrara, Italy.

Since 2021, he is with the Wireless Communication and Localization Networks Laboratory at the University of Ferrara. His research interests include network localization and navigation, and joint sensing and communication.

Mr. Torsoli serves as a reviewer for various IEEE journals and international conferences.

**Moe Z. Win** (Fellow, IEEE) is a Professor at the Massachusetts Institute of Technology (MIT) and the founding director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and with NASA Jet Propulsion Laboratory.

His research encompasses fundamental theories, algorithm design, and network experimentation for a broad range of real-world problems. His current research topics include ultra-wideband systems, network localization and navigation, network interference exploitation, and quantum information science. He has served the IEEE Communications Society as an Elected Member-at-Large on the Board of Governors, as elected Chair of the Radio Communications Committee, and as an IEEE Distinguished Lecturer. Over the last two decades, he held various editorial positions for IEEE journals and organized numerous international conferences. Recently, he has served on the SIAM Diversity Advisory Committee.

Dr. Win is an elected Fellow of the AAAS, the EURASIP, the IEEE, and the IET. He was honored with two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, co-authored with students and colleagues, have received several awards. Other recognitions include the IEEE Communications Society Edwin H. Armstrong Achievement Award (2016), the Cristoforo Colombo International Prize for Communications (2013), the Copernicus Fellowship (2011) and the *Laurea Honoris Causa* (2008) from the Università degli Studi di Ferrara, and the U.S. Presidential Early Career Award for Scientists and Engineers (2004). He is an ISI Highly Cited Researcher.

**Andrea Conti** (Fellow, IEEE) is a Professor and founding director of the Wireless Communication and Localization Networks Laboratory with the University of Ferrara, Italy. Prior to joining the University of Ferrara, he was with CNIT and with IEIIT-CNR.

In Summer 2001, he was with the Wireless Systems Research Department, AT&T Research Laboratories. Since 2003, he has been a frequent visitor to the Wireless Information and Network Sciences Laboratory at the Massachusetts Institute of Technology, Cambridge, MA, USA, where he presently holds the Research Affiliate appointment. His research interests involve theory and experimentation of wireless communication and localization systems. His current research topics include network localization and navigation, distributed sensing, adaptive diversity communications, and quantum information science.

Dr. Conti has served as editor for IEEE journals and chaired international conferences. He was elected Chair of the IEEE Communications Society's Radio Communications Technical Committee and is Co-Founder of the IEEE Quantum Communications & Information Technology Emerging Technical Subcommittee. He received the HTE Puskás Tivadar Medal, the IEEE Communications Society's Fred W. Ellersick Prize, and the IEEE Communications Society's Stephen O. Rice Prize in the field of Communications Theory. He is an elected Fellow of the IEEE and of the IET, and a member of Sigma Xi. He has been selected as an IEEE Distinguished Lecturer.