

Efficient Localization via Soft Information With Generic Sensing Measurements

Stefania Bartoletti^{ID}, *Member, IEEE*, Santiago Mazuelas^{ID}, *Senior Member, IEEE*,
Andrea Conti^{ID}, *Fellow, IEEE*, and Moe Z. Win^{ID}, *Fellow, IEEE*

Abstract—Accurate location awareness is essential for various context-based applications. This calls for efficient methodologies to collect, communicate and process position-dependent measurements, especially in situations with limited computational resources. The soft information (SI) approach has recently shown significant improvements in accuracy over conventional localization methods. By developing efficient SI-based techniques, it is possible to achieve higher precision also in case of stringent computational constraints. This paper proposes new SI-based localization techniques that utilize belief condensation and maximum entropy methods to reduce both communication burden and computational complexity. In addition, the techniques presented enable the use of generic sensing measurements, including those taking discrete and categorical values. Through two case studies involving time and angle measurements, we demonstrate how the proposed approach can significantly improve localization accuracy and computational efficiency.

Index Terms—Localization, soft information, belief condensation, machine learning, wireless networks.

I. INTRODUCTION

LOCATION awareness [1], [2], [3], [4], [5] is an essential component in next-generation networks [6], [7], [8], [9], [10] for a wide range of applications, including autonomy [11], [12], [13], assets tracking [14], [15], [16], crowdsensing [17], [18], [19], industrial IoT [20], [21], [22], and network optimization [23], [24], [25]. The design of localization systems and technologies needs to take into account

multiple aspects. The network architecture, localization protocols, signal processing, and quality of measurements play critical roles in terms of accuracy and coverage, as well as in terms of latency, availability, integrity, and overall system performance. The infrastructure complexity is also a crucial aspect, as localization capabilities need to be integrated within the existing infrastructure or the deployment of dedicated connections or devices might be required. In recent years, research has focused on integrated sensing, localization, and communication, particularly in the context of beyond 5G networks [26], [27], [28], [29], [30], [31]. While wireless localization uses radio signals to estimate the location of active devices, wireless sensing techniques use radio signals to detect events or changes in an environment, including the presence and location of passive targets. For example, radar-like sensing enables the detection and localization of unconnected (i.e., passive) targets by considering signal reflections from the target. As another example, sensing can be used to learn the wireless environment (e.g., multipath propagation, non-line-of-sight conditions) during the localization of active users. In this paper, we adopt the term sensing measurements in a broad sense to streamline the handling of position-dependent measurements for both passive targets and active users.

The main goal of localization systems is to estimate the target position (whether active users or passive objects) based on sensing measurements and contextual data. Sensing measurements are related to positional features such as range, angle, and signal power, and are obtained through wireless communication and signal processing among network nodes [32], [33], [34]. Contextual data, which can include digital maps, mobility models, and other relevant environmental information, may be known a priori or learned over time. Sensing measurements are influenced by various factors, including the wireless environment and the characteristics of the transmitter and receiver (e.g., antenna radiation patterns, receiver sensitivity).

The measurements at different nodes are first communicated over the wireless network, then processed and fused through a localization algorithm to obtain the position estimate [9], [35], [36], [37], [38], [39], [40]. For instance, fingerprinting, triangulation, or multilateration techniques are commonly used in both active and passive localization. Their implementation requires intra-node communication of single-value estimates (SVEs) of positional features (e.g., each measurement

Received 31 July 2024; revised 23 December 2024; accepted 15 January 2025. Date of publication 5 June 2025; date of current version 14 July 2025. The fundamental research described in this paper was supported, in part, by the European Union under ERC Grant 101078411, by the Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación under Grant PID2022-137063NB-I00, by the European Union under the NextGenerationEU Italian NRRP PE00000001 program RESTART, by the National Science Foundation under Grant CNS-2148251, by funds from federal agency and industry partners in the RINGS program, and by the Robert R. Taylor Professorship. The associate editor coordinating the review of this article and approving it for publication was C.-K. Wen. (*Corresponding author: Moe Z. Win.*)

Stefania Bartoletti is with the Department of Electronics Engineering and CNIT, University of Rome “Tor Vergata,” 00139 Rome, Italy (e-mail: stefania.bartoletti@uniroma2.it).

Santiago Mazuelas is with the Basque Center for Applied Mathematics (BCAM) and also the Ikerbasque-Basque Foundation for Science, 48009 Bilbao, Spain (e-mail: smazuelas@bcamath.org).

Andrea Conti is with the Department of Engineering and CNIT, University of Ferrara, 44122 Ferrara, Italy (e-mail: 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).

Digital Object Identifier 10.1109/TWC.2025.3540081

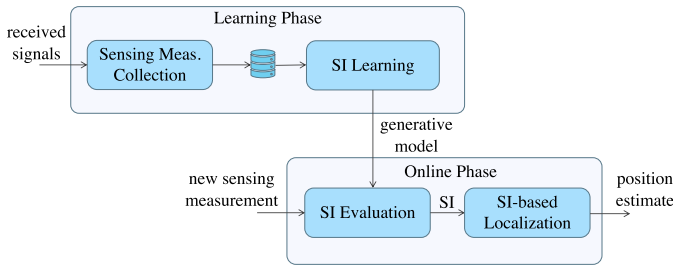


Fig. 1. Illustration of the main steps for SI-based localization.

corresponds to a single angle value or single range value). For example, 5G positioning has introduced various localization methods within the same technology framework. In this methods, the measurements can be collected either at UE-side or gNB-side, and the processing can be carried out either in the core network or locally at the radio access network (RAN). These aspects impact several localization performance indicators, including accuracy, availability, latency, and integrity.

In recent years, there has been a growing interest in the use of machine learning (ML) to learn the operating environment and channel conditions as well as to fuse different types of data, thus further enhancing the localization accuracy [1], [41], [42], [43]. In this context, the soft information (SI) framework has recently shown to significantly improve localization accuracy in multiple scenarios including those of beyond 5G networks [44], [45], [46], [47]. While classical algorithms rely on SVEs of positional features, SI-based techniques quantify the probability of positional features for any possible value. For SI-based localization, training data is collected at different nodes to learn generative models as depicted in Fig. 1. Such data-based learning approach is used to statistically characterize the sensing measurements.

In terms of accuracy, SI-based techniques excel over SVE-based techniques since they can extract all the available location information from measurements. Such richer information is encapsulated into the SI functions and enables the mitigation of wireless propagation impairments, as demonstrated in several studies [44], [45], [46]. Nevertheless, the usage of SI-based localization faces two distinct challenges. First, unlike SVE-based techniques, harnessing richer location information necessitates the evaluation of continuous functions, which are processed and communicated across various nodes within the network. In several operation conditions, there is a need to simplify both the SI learning process and operation phase to enhance the overall efficiency of the localization system.

Second, SI-based techniques have been exploited so far for handling continuous valued measurements such as waveform samples, time, and angle measurements. While most sensing measurements are commonly characterized as continuous random variables, it is worth noting that emerging sensing technologies frequently incorporate categorical or discrete measurements. One illustrative example is the measurement of discrete angles with a limited set of values as a result of beam sweeping approaches, which find application in wireless communications and localization as well as radar and optical sensing [48], [49], [50], [51], [52], [53], [54], [55]. The

growing use of categorical and discrete sensing data calls for adaptive learning approaches that can handle both continuous and discrete measurement types. To achieve this, generative models that seamlessly utilize discrete or categorical data are crucial.

This paper extends the SI approach to reduce communication and processing loads while seamlessly accommodating a variety of sensing measurements. To achieve this goal, we utilize two methods previously proposed in the existing literature but applied to other contexts, i.e., belief condensation (BC) and maximum entropy methods. The BC method has been utilized in [56], [57], and [58] to represent complex distributions in filtering and tracking applications where traditional techniques like Kalman filters and particle filters are inadequate. In this paper, we use the BC method to alleviate the burdens of communication and processing in SI-based localization. The maximum entropy method has been considered in the literature [59], [60], [61], [62], [63], [64], [65] for assessing event probabilities using distributions that maximize entropy within specified constraints. In this paper, we use maximum entropy models (MEMs) to extend the training capabilities and enable the effective handling of generic data in SI-based localization.

The key contributions of this paper can be summarized as in the following.

- We propose methods that use BC for obtaining multiple value estimates (MVEs) of positional features, leading to a more efficient communication and processing of the SI.
- We develop maximum entropy model (MEM)-based learning that can incorporate generic sensing measurements and rely on efficient convex optimization problems.
- We evaluate the effectiveness of the proposed approaches in terms of localization accuracy and computational complexity using continuous and discrete sensing measurements.

The performance of the proposed approaches is evaluated through sample-level simulations in two case studies: (i) time-based localization leveraging time difference-of-arrival (TDOA); and (ii) angle-based localization leveraging beam-sweeping for direction of arrival (DOA). The results demonstrate the advantages of the proposed methods in different operating conditions and for different localization algorithms.

The remainder of the paper is organized as follows: Sec. II describes the problem formulation and gives a review of the SI framework; Sec. III introduces the BC method for reducing the dimensionality of SI and thus leading to more efficient SI communication and processing; Sec. IV introduces the notion of MEM-based learning for SI-based localization; Sec. V evaluates the performance of the proposed techniques and compares it with existing SVE-based and SI-based techniques; and Sec. VI provides the main conclusion in the paper.

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 (RV) and its realization are denoted by x and x ; a random vector and its realization are denoted by $\mathbf{x}$ and $\mathbf{x}$, respectively. Sets and

random sets are denoted by upright sans serif and calligraphic font, respectively. For example, a random set and its realization are denoted by $\mathbf{X}$ and $\mathcal{X}$, respectively. The function $f_{\mathbf{x}}(x)$ and, for brevity when possible, $f(x)$ denote the probability distribution function (PDF) of a continuous RV $\mathbf{x}$; $\varphi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the PDF of a Gaussian random vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$; operator $\mathbb{E}_f\{\cdot\}$ denotes the expectation of the argument with respect to distribution f . For a matrix $\mathbf{A}$ and a vector $\mathbf{a}$, the transpose is denoted respectively by $\mathbf{A}^T$ and $\mathbf{a}^T$. Given two vectors $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{a} \preceq \mathbf{b}$ denotes that their i th components satisfy $a_i \leq b_i$ for $i = 1, 2, \dots, n$. The function $\mathbb{1}_{\{\mathcal{A}\}}(x) = 1$ if $x \in \mathcal{A}$ and 0 otherwise.

II. PROBLEM FORMULATION

In a localization problem, the goal is to estimate a state vector $\mathbf{x}$ using a set of n measurements $\{\mathbf{y}_i\}_{i=1}^n$, taken at a single node or at multiple nodes in a network. For example, nodes can perform multiple time measurements (e.g., round-trip-time or time-of-arrival) or angle measurements (e.g., angle-of-arrival or angle-of-departure) with respect to the target [66], [67], [68]. The state vector $\mathbf{x}$ is an array of dimension D with components that can include the position, velocity, and acceleration of a target. Each measurement $\mathbf{y}_i$ is an array of dimension M , with elements corresponding to a raw waveform or position-related data such as power, timing, and angle, which are usually extracted from the i th received waveform (either at the same or a different node). In general, the elements of $\mathbf{y}_i$ can have values that are continuous (e.g., waveform samples and power measurements from the analog front-end), discrete (e.g., angle or distance values with limited resolution due to hardware limitations), or categorical (e.g., target classification or antenna type).

The measurements $\mathbf{y}_i$ depend on the state vector $\mathbf{x}$ through a feature vector $\boldsymbol{\theta}_i$, which is a function of state vector $\mathbf{x}$, i.e., $\boldsymbol{\theta}_i(\mathbf{x})$. In this paper, we focus on the case where the positional feature takes scalar values.¹ For example, $\theta_i \in \mathbb{R}$ can be the distance between the transmitted and the i th receiver or can be the bearing angle between the transmitter and the i th receiver.

A. Localization Problem

By modeling the state vector as a deterministic unknown parameter and assuming that measurements are independent conditionally on $\mathbf{x}$, the maximum likelihood estimate (MLE) of the state vector is [2]

$$\begin{aligned} \hat{\mathbf{x}} &= \underset{\mathbf{x}}{\operatorname{argmax}} f(\{\mathbf{y}_i\}_{i=1}^n | \mathbf{x}) = \underset{\mathbf{x}}{\operatorname{argmax}} \prod_{i=1}^n f(\mathbf{y}_i | \mathbf{x}) \\ &= \underset{\mathbf{x}}{\operatorname{argmax}} \prod_{i=1}^n f(\mathbf{y}_i | \theta_i). \end{aligned} \quad (1)$$

The localization process aims to solve the optimization problem in (1), which requires the evaluation of $f(\mathbf{y}_i | \theta_i)$ for each $i = 1, 2, \dots, n$.

¹In the sequel, when there is no ambiguity, we will omit $\theta_i(\mathbf{x})$'s explicit dependence on $\mathbf{x}$ and simply write θ_i for notational convenience.

The conditional distribution $f(\mathbf{y}_i | \theta_i)$ is unknown in general and conventional methods use simple models to approximate the relationship between the measurements and the positional feature. Usually, such relationship is described by using a SVE $\hat{\boldsymbol{\theta}}_i(\mathbf{y}_i)$ so that $f(\mathbf{y}_i | \theta_i) \propto f(\hat{\boldsymbol{\theta}}_i(\mathbf{y}_i) | \theta_i)$ which is modeled by a simple distribution such a Gaussian with mean θ_i . In these cases, the MLE becomes the least square estimate (LSE) for cases with same variance or the weighted least square estimate (WLSE) for cases with different variances. In particular, the LSE leads to

$$\hat{\mathbf{x}} = \min_{\mathbf{x}} \sum_{i=1}^n |\theta_i(\mathbf{x}) - \hat{\boldsymbol{\theta}}_i(\mathbf{y}_i)|^l \quad (2)$$

where $\theta_i(\mathbf{x})$ is the positional feature value corresponding to $\mathbf{x}$ and $l = 2$. It is recommended to use WLSE for heterogeneous and independent measurements (e.g., range and angle data from different nodes) to account for different variances and ensure proper normalization [71]. Minimum norm optimization in (2) can be efficiently solved using conventional methods and readily available solvers. The use of the L1-norm (i.e., $l = 1$) instead of the L2-norm (i.e., $l = 2$) can improve the performance in some scenarios, especially in the presence of outliers [69], [70].

The framework of SI-based localization has been recently proposed in [44], [45], and [46] as a learning-based solution to the limitations of SVE-based localization. SI can incorporate all the positional information of the target provided by measurements. Indeed, SI is obtained by learning the joint distribution of positional features and measurements from off-line measurements. In particular, the SI of measurement $\mathbf{y}_i$ for positional feature θ_i , denoted as $\mathcal{L}_{\mathbf{y}_i}(\theta_i)$, is proportional to $f(\mathbf{y}_i | \theta_i)$ in the absence of prior knowledge for the positional feature θ_i (using a constant reference prior [45]). If prior knowledge or contextual information for the positional feature θ_i is available (e.g., digital maps, mobility models), then θ_i is treated as a RV according to a Bayesian approach, and we have that $\mathcal{L}_{\mathbf{y}_i}(\theta_i) \propto f(\theta_i, \mathbf{y}_i) / f(\theta_i)$. Using SI, the MLE in (1) becomes

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmax}} \prod_{i=1}^n \mathcal{L}_{\mathbf{y}_i}(\theta_i). \quad (3)$$

Note that (3) can be used also in case of heterogeneous and multimodal measurements, such as combining range and angle data from different nodes [44], [46].

B. Existing Approaches for SI Learning and Processing

As a general approach, SI is obtained in two phases illustrated in Fig. 1. During the training phase, a dataset of S training samples is collected as $\{\theta_k, \mathbf{y}_k\}_{k=1}^S$, where θ_k is the ground truth value of the positional feature. The training samples are used to learn the generative model for the sensing measurement, i.e., the joint distribution $f(\theta, \mathbf{y})$.

Existing approaches for SI learning utilize a Gaussian mixture model (GMM) to approximate the joint distribution $f(\theta, \mathbf{y})$ by leveraging the dataset of training samples

$\{\theta_k, \mathbf{y}_k\}_{k=1}^S$ [72], [73]. In particular, a mixture of K Gaussians is used to approximate the distribution $f(\theta, \mathbf{y})$ as

$$\hat{f}(\theta, \mathbf{y}) = \sum_{i=1}^K \alpha_i \varphi((\theta, \mathbf{y}); \mathbf{m}_i, \Sigma_i) \quad (4)$$

where $\varphi(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes the PDF of a Gaussian random vector $\mathbf{x}$ with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Such an approximation $\hat{f}(\theta, \mathbf{y})$ can be obtained using the expectation maximization (EM) algorithm as detailed in [45].

During the online phase, the SI function $\mathcal{L}_{\mathbf{y}'}(\theta)$ for a new measurement $\mathbf{y}'$ is obtained by plugging in such measurement in the learned generative model, i.e., $\mathcal{L}_{\mathbf{y}'}(\theta) \propto \hat{f}(\theta, \mathbf{y}')$. Likewise, the SI function corresponding to the data fusion of p measurements $\mathbf{y}' = [\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_p]$ can be obtained as

$$\mathcal{L}_{\mathbf{y}'}(\theta) \propto \hat{f}(\theta, \mathbf{y}'_1) \hat{f}(\theta, \mathbf{y}'_2) \dots \hat{f}(\theta, \mathbf{y}'_p) \quad (5)$$

as long as $\mathbf{y}'_1, \mathbf{y}'_2, \dots, \mathbf{y}'_p$ are conditionally independent given θ [44].

During the localization phase, multiple SIs $\mathcal{L}_{\mathbf{y}_1}(\theta_1), \mathcal{L}_{\mathbf{y}_2}(\theta_2), \dots, \mathcal{L}_{\mathbf{y}_n}(\theta_n)$ are collected from n different sources (e.g., anchors), and the positional state can be estimated using (3). Existing methods for SI-based localization address such optimization by evaluating the SIs functions in a grid of possible values for $\mathbf{x}$ and selecting $\hat{\mathbf{x}}$ as the grid element that maximizes the values of (3) [44], [45].

C. Challenges of Existing Approaches

SI-based localization provides a significant performance improvement with respect to SVE-based localization, especially in terms of robustness to wireless propagation conditions and hardware limitations [44], [45], [46]. In the following, we describe the main challenges of SI-based localization in terms of processing and communication efficiency as well as the generality of the measurements that are considered.

1) *Communication and Processing Challenges*: In terms of position estimation, solving the optimization problem (3) through an exhaustive grid-based search leads to a complexity that rapidly increases with the grid resolution. Specifically, in terms of number of grid levels per dimension such complexity is quadratic or cubic for 2D or 3D localization, respectively. Therefore, grid-based techniques are challenging to implement in large-scale environments.

In terms of communication costs, the inter-node communication required to exchange and process SIs can be also highly inefficient. To implement SI-based localization in a network of nodes, either the learned model parameters and measurements can be shared in the network or the SI can be obtained locally for each measurement and communicated to other nodes. As an example, when $\mathbf{y}_i$ is the vector of received signal samples at the i th receiver in a network, either the samples $\mathbf{y}_i$ with the learned model $f(\theta_i, \mathbf{y}_i)$ or $\mathcal{L}_{\mathbf{y}_i}(\theta_i)$ need to be communicated to a central node for fusion. The former approach can result in unaffordable communication complexities in cases with high-dimensional measurements or model parameters. In addition, such approaches may undermine the privacy of different agents since the learned model can contain user-specific information. In the latter approach, the communication load can be considerably high since the SI

is a general continuous function of θ_i determined by a possibly large number of parameters.

2) *Learning Challenges for Generic Sensing Measurements*: The importance of generative model estimation according to (4) is evident across various scenarios, benefiting from the adaptability of mixtures of Gaussians in accurately approximating general continuous distributions. However, the use of mixtures of Gaussians results in challenging learning processes [72], [73]. For instance, maximizing the likelihood over mixtures of Gaussians is a non-convex optimization problem and existing methods like the EM algorithm only find a local optimum and rely on the initialization used.

As another limitation of the usage of mixtures of Gaussians, there are cases where modeling based on Gaussian distributions is inappropriate, particularly in describing discrete or categorical values. The use of beam scanning and pairing for sensing, localization, and communication, often involves discrete angle measurements [52], [53], [54], [55]. In most cases, angle estimates are constrained to a very limited set of values due to hardware limitations, i.e., the number of antennas, as well as latency constraints on the beam sweeping operation [72], [73], [74], [75]. To effectively incorporate these types of measurements for SI-based localization, generative models capable of handling discrete data are required. However, Gaussian distributions or mixtures of them are inappropriate for modeling discrete RVs.

III. SI-BASED LOCALIZATION VIA BELIEF CONDENSATION

This section describes an approach leveraging BC method for efficient SI-based localization. The presented approach is described in Fig. 2; the online phase of the localization process is developed in two steps: (i) different nodes in the network process new measurements $\mathbf{y}'$ (e.g., time or angle measurements) to obtain multiple estimates of the parameter θ , i.e., a finite set Θ_{BC} ; (ii) a centralized processor uses all the Θ_{BC} to obtain the estimated state vector $\hat{\mathbf{x}}$. The BC method is used to obtain Θ_{BC} by approximating the SI $\mathcal{L}_{\mathbf{y}}(\theta)$.

A. SI Approximation via Belief Condensation

As illustrated in Fig. 2, the BC step takes as input the new measurement $\mathbf{y}'$ and the learned generative model $f(\theta, \mathbf{y})$ and provides as output a set of MVEs, denoted by Θ_{BC} . For example, if $\mathbf{y}'$ is a distance measurement and θ is the true distance, the output Θ_{BC} contains feasible values for the true distance. The main idea is to approximate the SI continuous function with a finite set of values, namely MVEs, that can be processed more efficiently. In the following we show how to obtain such approximation by minimizing the Kolmogorov-Smirnov (KS) distance. The KS distance is a measure of the difference between two probability distributions, which is calculated as the maximum absolute difference between the cumulative distribution functions (CDFs) [76]. The KS distance is non-parametric and does not make any assumptions about the shape or parameters of the underlying distributions being compared. The SI function $\mathcal{L}_{\mathbf{y}}(\theta)$ is then approximated by a function supported in r values $\Theta_{\text{BC}} = \{\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r\}$,

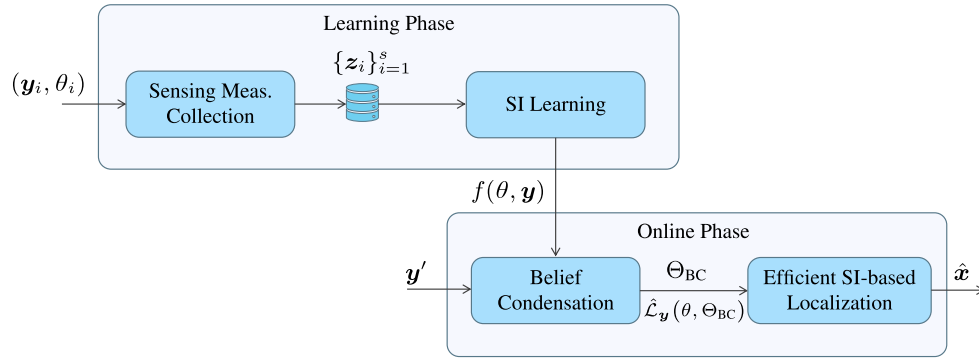


Fig. 2. Illustration of the main steps for efficient SI-based localization based on BC.

which correspond to the MVEs. The set Θ_{BC} is chosen as the one that minimizes the KS distance with respect to the normalized SI. In particular, we define the CDF induced by a SI as

$$F_{\mathcal{L}}(\theta) = \frac{\int_{-\infty}^{\theta} \mathcal{L}_{\mathbf{y}}(\varsigma) d\varsigma}{\int_{-\infty}^{+\infty} \mathcal{L}_{\mathbf{y}}(\varsigma) d\varsigma}. \quad (6)$$

The following theorem shows how the KS distance can be minimized by using a piecewise CDF obtained from the quantiles of $F_{\mathcal{L}}(\theta)$.

Theorem 1: Let $F_{\mathcal{L}}$ be the CDF induced by a SI $\mathcal{L}_{\mathbf{y}}(\theta)$. If $\mathcal{F}_r$ is the set of CDFs with r support points in the real line, and $\hat{F}_{\mathcal{L}}(\theta) = \frac{1}{r} \sum_{i=1}^r \mathbb{1}\{\theta \leq \hat{\theta}_i\} \in \mathcal{F}_r$ with $F_{\mathcal{L}}(\hat{\theta}_i) = \frac{2i-1}{2r}$, then

$$\hat{F}_{\mathcal{L}}(\theta) \in \arg \min_{\hat{F} \in \mathcal{F}_r} \left(\sup_{\theta \in \mathbb{R}} |F_{\mathcal{L}}(\theta) - \hat{F}(\theta)| \right)$$

and

$$\sup_{\theta \in \mathbb{R}} |F_{\mathcal{L}}(\theta) - \hat{F}_{\mathcal{L}}(\theta)| = \frac{1}{2r}.$$

The theorem above follows from the results in Proposition 3 and Proposition 4 in [56] for the addressed setting, i.e., by substituting $F(\mathbf{x})$, $G_{\mathbf{x}}$, and m in [56] with $F_{\mathcal{L}}(\theta)$, $\hat{F}_{\mathcal{L}}(\theta)$, and r , respectively. The above result shows that the uniform distribution supported on the $\frac{2i-1}{2r}$ -quantiles for $i = 1, 2, \dots, r$ of the normalized $\mathcal{L}_{\mathbf{y}}(\theta)$ is the probability distribution, having smallest KS distance to the normalized $\mathcal{L}_{\mathbf{y}}(\theta)$. The resulting support points $\hat{\theta}_i$ are referred to as the MVEs for θ .

Selecting the r MVEs according to Theorem 1 brings multiple advantages, as described in the following.

- The MVEs $\{\hat{\theta}_i\}_{i=1}^r$ that approximate the SI have a clear interpretation. Specifically, they correspond to equally likely values for θ when the KS-closest distribution to the SI is considered.
- Storing, communicating, and processing r values is much more efficient than handling the entire function $\mathcal{L}_{\mathbf{y}}(\theta)$.
- SI-based localization leveraging MVEs can rely on simpler algorithms similar to those used for SVEs.
- The MVEs in Θ_{BC} can be easily calculated using Theorem 1 through Algorithm 1.

Algorithm 1 shows the pseudocode for approximating the SI $\mathcal{L}_{\mathbf{y}}(\theta)$ in the interval $(\theta_{\min}, \theta_{\max})$. For a discretization of

Algorithm 1 SI Approximation Through BC

Input: SI $\mathcal{L}_{\mathbf{y}}(\theta)$, $\theta_{\min}$, $\theta_{\max}$, $r > 0$, $h \geq r$

Output: MVEs $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_r$

$\Delta \leftarrow (\theta_{\max} - \theta_{\min})/h$

$\vartheta_1 \leftarrow \theta_{\min}$

$q \leftarrow 0$

$i \leftarrow 1$

for $j = 2, \dots, h$ **do**

$\vartheta_j \leftarrow \vartheta_{j-1} + \Delta$

$c_j \leftarrow \mathcal{L}_{\mathbf{y}}(\vartheta_j)$

$F_j \leftarrow \Delta \sum_{k=1}^j c_k$

end for

for $j = 1, 2, \dots, h$ **do**

if $(q - F_j/F_h) > 0$ **then**

$\hat{\theta}_i \leftarrow \vartheta_j$

else

if $q > (F_j - F_{j-1})/2F_h$ **then**

$\hat{\theta}_i \leftarrow \vartheta_j$

end if

if $i = r$ **then break**

else

$i \leftarrow i + 1$

$q \leftarrow q + 1/r$

end if

end if

end for

$(\theta_{\min}, \theta_{\max})$ with h levels, such algorithm has a computational complexity $\mathcal{O}(h)$ and can be easily implemented using h evaluations of the SI. Note that finding the quantiles corresponding to the BC support points requires only $\mathcal{O}(h)$ operations, as shown in the second loop in Algorithm 1. Specifically, the quantiles can be obtained sequentially with a single pass through the sequence of CDF values, given the inherently non-decreasing nature of this sequence.

Fig. 3 illustrates an example of BC for soft range information, i.e., y is a range measurement for the distance feature θ . In the example, the range measurements are drawn from a RV that follows a multi-modal distribution. During the training phase, 100 samples of range measurement and ground truth values are collected and used to estimate the generative model through GMM fitting. The BC corresponding to the SI $\mathcal{L}_{\mathbf{y}}(\theta)$ is obtained with $r = 10$ and $h = 1000$, following

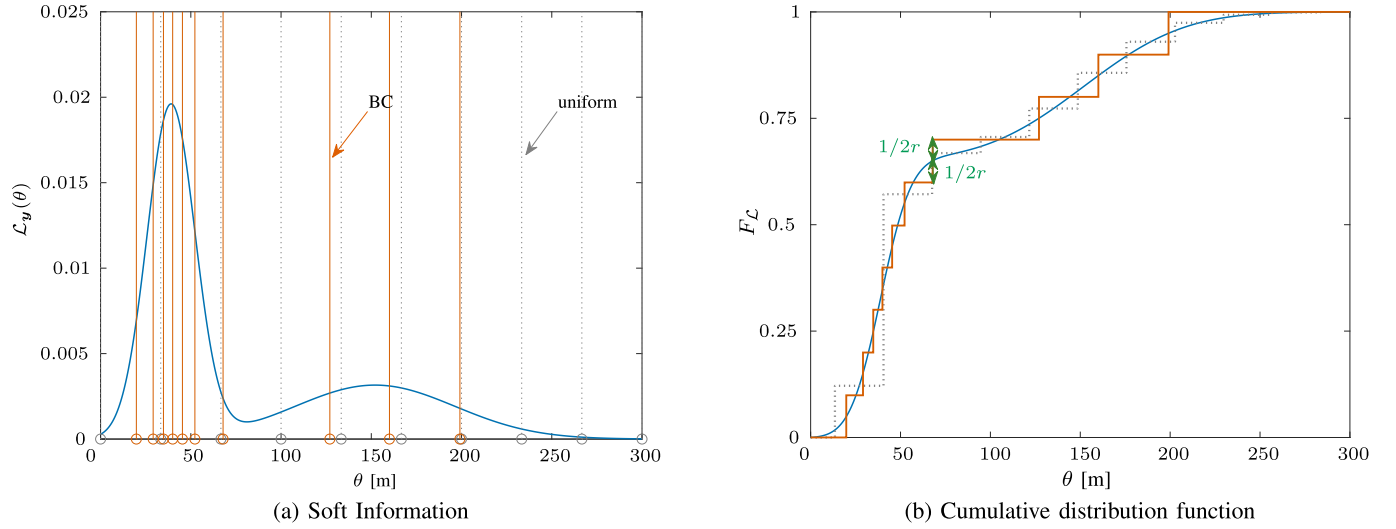


Fig. 3. Illustration of BC of soft range information in an example scenario. The blue solid lines represent the SI (a) and its CDF (b). The red solid lines represent the discrete set of range values (a) and the piecewise approximated CDF (b) obtained through BC. The gray dotted lines represent the discrete set of range values (a) and the piecewise approximated CDF (b) obtained with uniform sampling of the SI.

Algorithm 1. In particular, Fig. 3a shows the continuous SI (blue curve) and compares it to the MVEs in Θ_{BC} obtained through BC. A further comparison is made with the case where uniform sampling is used. It can be noted that the MVEs obtained with BC are able to capture the multimodality of the distribution, with more samples around the most likely values. Differently, the MVEs obtained with the uniform sampling have many samples in less likely values. Fig. 3b shows the corresponding CDF and compares it with the empirical CDF of the MVEs for both the BC and uniform sampling. Such figure also shows that the proposed approximation achieves the smallest KS distance, since the maximum difference between the continuous and discrete CDF is $1/2r$. Fig. 4 illustrates the KS distance between $\hat{\mathcal{L}}_y(\theta)$ and $\mathcal{L}_y(\theta)$ as a function of the number of MVEs r . The results are compared using the two methods of uniform sampling and BC. It is evident that the KS distance is lower for the BC, even with high values of r . This indicates that BC is more effective to approximate $\mathcal{L}_y(\theta)$.

B. SI-Based Localization via MVEs

As illustrated in Fig. 2, the localization step takes as input the MVEs Θ_{BC} representing multiple feasible values for the feature θ (e.g., true distance), and provides as output the estimated position $\hat{x}$. Let $\mathcal{L}_{y_i}(\theta_i)$ for $i = 1, 2, \dots, n$ be the SIs corresponding to measurements $\{y_i\}_{i=1}^n$ and $\Theta_{BC,i} = \{\hat{\theta}_{i,1}, \hat{\theta}_{i,2}, \dots, \hat{\theta}_{i,r}\}$ be the MVEs obtained by Algorithm 1. We now show how the use of $\Theta_{BC,i}$ can significantly improve the communication and processing efficiency.

1) *Norm Minimization*: MVE-based localization can be directly achieved by extending (2) as

$$\hat{x} = \min_x \sum_{i=1}^n \sum_{j=1}^r |\theta_i(x) - \hat{\theta}_{i,j}|^l. \quad (7)$$

This optimization problem can be solved similarly to that in (2), i.e., the computational complexity is similar to that

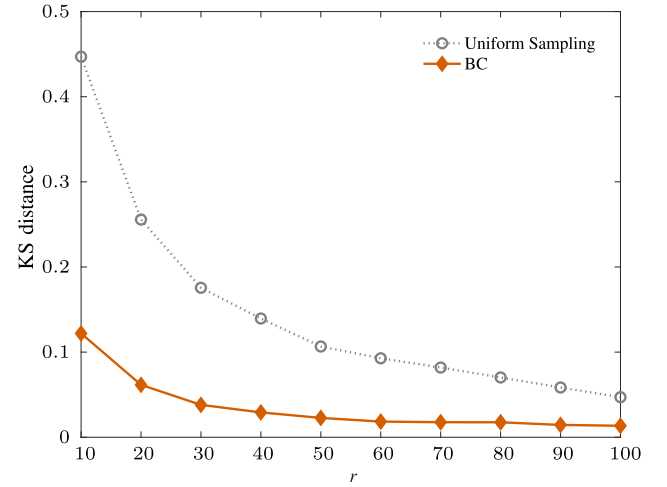


Fig. 4. KS distance between $\hat{\mathcal{L}}_y(\theta)$ and $\mathcal{L}_y(\theta)$ as a function of the number of MVEs r using uniform sampling and belief condensation.

of SVE-based techniques. Specifically, the optimization problem (7) is as tractable as that for LS in (2) and their main difference is that the objective function (7) has nr terms instead of n .

2) *Statistical Interpolation of MVEs*: An alternative approach is proposed here for scenarios where it is necessary to reduce the inter-node communication load, yet there are less stringent requirements in terms of processing. In this case, the MVEs obtained through BC can be statistically interpolated to reconstruct the SI in a continuous form. In particular, the approximate SI $\hat{\mathcal{L}}_y(\theta)$ can be obtained by statistically interpolating the MVEs Θ_{BC} using histogram-based density estimation as

$$\hat{\mathcal{L}}_y(\theta) = \frac{1}{rh} \sum_{q=1}^r K\left(\frac{\theta - \hat{\theta}_q}{h}\right) \quad (8)$$

where $K(x) = \mathbb{1}_{\{(-1,1]\}}(x)$ and h defines the width of each histogram interval.

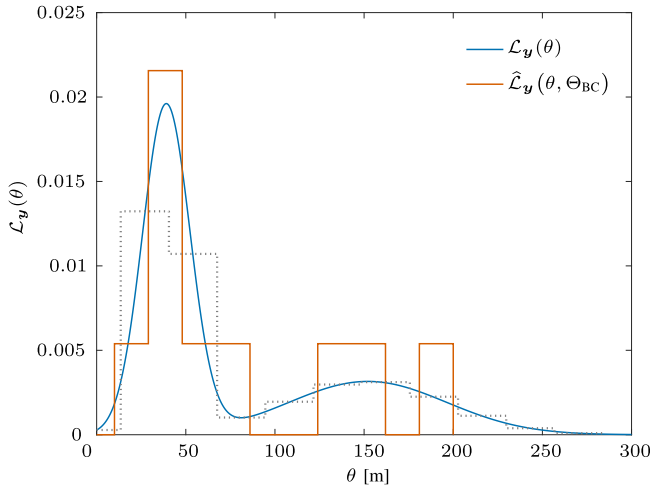


Fig. 5. Illustration of SI with BC and statistical interpolation in the example scenario. The blue solid line represents the continuous SI function $\mathcal{L}_y(\theta)$. The gray dotted line represents an approximate SI with uniform sampling. The red solid line represents the approximate SI with BC and statistical interpolation, i.e. $\hat{\mathcal{L}}_y(\theta)$.

Fig. 5 shows the effect of statistical interpolation and compares the original SI and the approximate SI with BC and statistical interpolation in the same setting of Fig. 3. It can be seen that BC together with statistical interpolation enables the reconstruction of the main modes of the original distribution. The figure also compares two SI approximations obtained via BC and uniform sampling. It can be seen that the approximate SI obtained via uniform sampling is poorer than that obtained via BC, especially for the most likely values of θ .

C. Discussion

The methods presented above can enable to reduce the communication overhead in SI-based localization while maintaining the benefits of using continuous functions representing the SIs. For example, different nodes in a heterogeneous network might have varying requirements in terms computational capabilities and resources, so that some of them can communicate using SI functions while others use MVEs. By statistically interpolating MVEs to obtain the approximate SI, we can exploit all available information and apply SI-based localization algorithms, thereby accommodating the diverse capabilities of the nodes and enhancing the overall accuracy and efficiency of the network.

The proposed BC method reduces the communication requirements and minimizes the KS distance to the original distribution. For SI-based techniques based on mixtures of Gaussians, other methods can also be used to approximate SI, such as those that involve pruning and merging components [77], [78], [79], [80]. However, these methods are designed to simplify mixtures of Gaussians by approximating various distance metrics with unknown errors and require complex procedures. On the other hand, the approximation based on BC method finds the closest distribution with high efficiency and can utilize a straightforward least squares algorithm for the efficient processing of MVEs, making it more advantageous overall.

The presented methodology for SI-based localization based on BC can result in efficient localization systems that can make use of common algorithms for SVE-based localization. The main benefits of the methods proposed in terms of communication and processing complexity are as follows.

- General SI functions can be accurately represented using a reduced number r of support points. As described above, the probabilistic error due to such approximation decreases at a rate $\mathcal{O}(1/r)$ for general SI functions. The experimental results in Section V show that even $r = 10$ support points can be enough to attain the benefits of SI. On the other hand, parametric SI-based techniques based on mixtures of Gaussians may require to use a large number of parameters to represent SI functions. For instance, if an agent combines P measurements with individual SIs given by mixtures of K components, the resulting SI would be given by a mixture with K^P components resulting from the multiplication of the P mixtures.
- Conventional localization algorithms can be used for SI-based localization. The r MVEs can be treated as multiple SVEs. This enables to reap the benefits of SI using conventional algorithms for SVE-based localization such as LS.

IV. SI VIA MAXIMUM ENTROPY

This section presents how the maximum entropy method can be applied to SI-based localization. In particular, MEMs are used to learn the generative model $f(\theta, \mathbf{y})$ and can be applied to handle sensing measurements of generic types, including those that are discrete or categorical. Examples of discrete variables are angles and range values with a finite resolution due to hardware limitations. Example of categorical measurements include target classification or antenna type, and any variable that falls in an unsorted category.

A. Maximum Entropy Method for SI Learning

The maximum entropy method advocates employing a distribution that maximizes entropy among those adhering to specific expectations' constraints [59], [60], [61]. In the following derivations, we apply this method to learn the joint distribution, denoted as $f(\theta, \mathbf{y})$, which maximizes Shannon entropy within the set of distributions satisfying a predefined set of expectations' constraints derived from training samples. This approach does not rely on the adequacy of modeling assumptions, such as Gaussian distributions, and only requires to estimate expectations using training data.

Let the random vector $\mathbf{z} = (\theta, \mathbf{y})$ be formed by pairs of positional features and their corresponding observations. We use the training samples $\{\mathbf{z}_i\}_{i=1}^s = \{\theta_i, \mathbf{y}_i\}_{i=1}^s$ to calculate constraints for the expectation of a function $\Phi(\mathbf{z}) \in \mathbb{R}^m$ defined on the support $\mathcal{Z}$ of the random vector $\mathbf{z}$, i.e., $\Phi : \mathcal{Z} \rightarrow \mathbb{R}^m$. Such functions Φ are usually known as feature mappings in the field of ML and are commonly formed by polynomials and threshold functions [62], [63], [64], [65]. In the beamsweeping example, $\mathbf{z} = (\theta, \mathbf{y})$ can be formed by pairs of angles and their corresponding observations. Training samples $\{\mathbf{z}_i\}_{i=1}^s = \{\theta_i, \mathbf{y}_i\}_{i=1}^s$ are composed by multiple

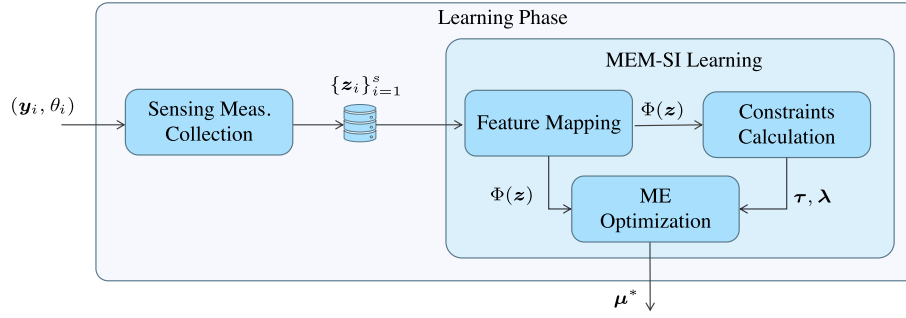


Fig. 6. Illustration of the main steps for MEM-based learning phase.

angle measurement obtained with respect to different angles during the training phase.

Consider an interval estimate for the expectation of the feature mapping Φ in the form $[\tau - \lambda, \tau + \lambda]$. The most common way to get such an estimate is using sample means and standard deviations as

$$\begin{aligned}\tau_j &= \frac{1}{s} \sum_{i=1}^s \Phi_j(z_i) \\ \lambda_j &= \frac{1}{s} \sqrt{\sum_{i=1}^s (\Phi_j(z_i) - \tau_j)^2}\end{aligned}\quad (9)$$

where τ_j and λ_j are the j th elements of τ and λ , respectively, with $j = 1, 2, \dots, m$. Then, we choose the distribution $f(z)$ that maximizes the entropy among those that have expectations aligning with the data-based estimates. Specifically, the generative model determined by MEM is the solution to the optimization problem

$$\begin{aligned}\underset{f}{\text{maximize}} \quad & -\mathbb{E}_f\{\log f(z)\} \\ \text{subject to} \quad & \tau - \lambda \preceq \mathbb{E}_f\{\Phi(z)\} \preceq \tau + \lambda\end{aligned}\quad (10)$$

that is a convex optimization problem since Shannon entropy is concave and the constraints above are linear on $f(z)$. Using convex duality, it can be shown (see, e.g., [81], [82], [83], [84]) that the solution of such optimization problem is

$$\hat{f}_{\text{MEM}}(z) = \exp\{\Phi(z)^T \mu^* + \nu^*\} \quad (11)$$

where $\mu^* \in \mathbb{R}^m$ is the vector of parameters solving the convex optimization problem

$$\underset{\mu}{\text{minimize}} \quad -\tau^T \mu + \log \int_{\mathcal{Z}} \exp\{\Phi(z)^T \mu\} d\eta(z) + \sum_{i=1}^m \lambda_i |\mu_i| \quad (12)$$

and

$$\nu^* = -\log \int_{\mathcal{Z}} \exp\{\Phi(z)^T \mu^*\} d\eta(z)$$

where $\eta(z)$ is the base measure in $\mathcal{Z}$. As ν^* results in a proportionality constant for $\hat{f}_{\text{MEM}}(z)$, its calculation can be avoided.

Highly efficient methods have been developed to solve the convex optimization problem in (12) [63]. The integral in the log-partition function in (12) must be approximated in practice. Various approaches for such approximations exist, including

uniform grids, Monte Carlo methods [82], and Gaussian quadrature [83]. For instance, using a Monte Carlo approach, the log-partition function in (12) can be approximated as

$$\log \int_{\mathcal{Z}} \exp\{\Phi(z)^T \mu\} d\eta(z) \approx \log \frac{1}{|\tilde{\mathcal{Z}}|} \sum_{z \in \tilde{\mathcal{Z}}} \exp\{\Phi(z)^T \mu\} \quad (13)$$

where the sample space $\tilde{\mathcal{Z}}$ is given by a sufficiently large number of samples.

Fig. 6 illustrates the main steps for SI learning through MEMs. The constraint calculation block defines the expectation constraints obtained from the training samples and (10). Then, the maximum entropy optimization problem is solved according to (12), thus obtaining the parameters μ^* and ν^* . Finally, the generative model is determined as $\hat{f}_{\text{MEM}}(z)$ in (11). The online phase and the determination of MVEs shown above can be carried out as described in Sec. III since those steps are independent of the method used to learn the SIs.

Note that the generative model derived in (11) via the maximum entropy method does not depend on any assumptions about the underlying data or its distribution, as it only requires generic training data. Consequently, the MEM is particularly well-suited for managing discrete and categorical data, which are often inadequately handled by existing SI learning methods that rely on Gaussian distributions or their mixtures. The numerical results in Sec. V showcase that the maximum entropy method empowers the utilization of the SI framework with discrete measurements, further enhancing localization performance in comparison to SVE-based techniques.

In the following subsections, we explore various feature mapping choices commonly used in other maximum entropy methods and introduce a novel approach tailored for measurements acquired in scenarios influenced by multimodal propagation conditions of wireless signals.

B. Feature Mappings for MEMs

Here we briefly introduce low-complexity feature mappings Φ that are used in existing applications of MEMs for generative model estimation [59], [60], [61], [62], [63], [64], [65]. These mappings are not developed for localization purposes but are commonly used in other types of systems, such as language modeling, among others. Starting from these basic features, we will then provide clustering-based features for SI-based localization.

1) *Extraction of Elementary Features*: in a first step, the original values in $\mathcal{Z}$, $\mathbf{z} = (\theta, \mathbf{y})$ of dimension F are transformed to elementary features $\mathbf{z}^{(1)}$ with dimensionality L_1 . For example, the l -th component of the feature mapping can be given by the j -th component of $\mathbf{z}$ as $z_l^{(1)} = z_j$ or can be a boolean value depending on the value of z_j as $z_l^{(1)} = \mathbb{1}_{\{z_j \in \mathcal{R}_l\}}$.

2) *Products of Elementary Features*: in a second step, elementary features $\mathbf{z}^{(1)}$ are transformed to product features $\mathbf{z}^{(2)}$ with dimensionality L_2 . Specifically, let $j_1, j_2, \dots, j_k \in \{1, 2, \dots, L_1\}$, the corresponding product feature of order k is given by $z_l^{(2)} = z_{j_1}^{(1)} \cdot z_{j_2}^{(1)} \cdot \dots \cdot z_{j_k}^{(1)}$. This process can be repeated to obtain multiple polynomials of elementary features. The feature mapping is then given by $\Phi(\mathbf{z}) = \Phi_1(\mathbf{z}), \Phi_2(\mathbf{z}), \dots, \Phi_m(\mathbf{z})$ where for $l = 1, 2, \dots, m$, $\Phi_l(\mathbf{z}) = z_l$ corresponds to one of the features described above.

3) *Dimensionality Reduction*: in this initialization step, the original values in $\mathcal{Z}$, $\mathbf{z} = (\theta, \mathbf{y})$ with dimensionality F are transformed to $\Phi^{(0)}(\mathbf{z}) = \mathbf{z}^{(0)}$ with dimensionality $L_0 \ll F$. Specifically, given a function $\psi(\cdot)$ that reduces dimensionality of measurements $\mathbf{y}$, the original values $\mathbf{z}$ are transformed as $\Phi^{(0)}(\mathbf{z}) = \mathbf{z}^{(0)} = (\theta, \psi(\mathbf{y}))$.

In common applications of maximum entropy, the two steps for the extraction of elementary features and their product are often applied to the vector $\mathbf{z}^{(0)} = \Phi^{(0)}(\mathbf{z})$ rather than directly to $\mathbf{z}$.

C. Clustering-Based Features for SI-Based Localization

We introduce a type of feature based on clustering. Such an approach is used to design MEM features that can model the prevalent multimodality of wireless signals in challenging environments characterized by multipath and non-line-of-sight (NLOS) conditions. By leveraging clustering techniques, we can effectively model this multimodality, thus distinguishing and grouping measurements (e.g., range, angle or power measurement) based on the propagation conditions that correspond to line-of-sight (LOS) and NLOS or are affected by multipath fading. The proposed clustering approach can help to define a feature mapping that is suitable for modeling measurements in wireless sensing and localization scenarios.

We leverage training data to identify m clusters corresponding to the values of $\mathbf{z}$ or any pre-extracted feature. Efficient acquisition of these clusters can be achieved using algorithms such as K-means [85]. For each cluster $j = 1, 2, \dots, m$, let $\mathbf{c}_j$ be its centroid (e.g., the average of samples in the j th cluster). For any $\mathbf{z} \in \mathcal{Z}$, let ξ_j be the squared distance to the j th centroid, that is

$$\xi_j = (\mathbf{z} - \mathbf{c}_j)^T (\mathbf{z} - \mathbf{c}_j)$$

for $j = 1, 2, \dots, m$ and $j(\mathbf{z})$ be the closest centroid to $\mathbf{z}$, that is

$$j(\mathbf{z}) = \underset{j=1,2,\dots,m}{\operatorname{argmin}} \xi_j$$

then the feature mapping at $\mathbf{z}$ is $\Phi(\mathbf{z}) \in \mathbb{R}^m$, with q th component

$$\phi_q(\mathbf{z}) = \mathbb{1}_{\{q\}}(j(\mathbf{z})) \xi_q \text{ for } q \in \{1, 2, \dots, m\}. \quad (14)$$

The selection of the feature mapping plays a pivotal role in determining the MEM since it encapsulates the attributes of positional features and sensing measurements estimated from training data. The performance of the BC and MEM proposed in Sec. III and IV will be evaluated in the following section both in terms of complexity as well as localization accuracy.

V. CASE STUDIES

This section presents results in two case studies: (i) TDOA-based localization and (ii) DOA-based localization via beam scanning. The first case study shows the efficiency provided by the usage of SI-based MVEs obtained through BC. The second case study shows how MEM-based SI can be used to handle discrete measurements. In particular, each measurement is the angle estimated through beam sweeping with low angular resolution. In both case studies, results obtained with BC and maximum entropy methods will be compared to the results obtained using classical SVE-based localization and using existing methods for SI.

A. Case Study 1: TDOA-Based Localization

The first case study takes as example application TDOA-based cellular localization, which is a method used to determine the location of a user equipment (UE) in a cellular network. It relies on measuring the reference signal time difference (RSTD) between neighboring gNodeB (gNB) and a reference gNB, where gNBs are the 5G New Radio (NR) base stations.

1) *Simulation Settings*: We consider a squared scenario with 500m edge length, with four gNBs deployed at the corners. The reference gNB is the one at the bottom left corner. For each random instantiation, the UE is in a random position $\mathbf{x}$, uniformly distributed within the square, and collects the measurement $\mathbf{y}_i$ which is the i th RSTD measurement, with $i = 1, 2, 3$. The positional feature is $\theta_i(\mathbf{x})$, i.e., the true RSTD value for a UE being at $\mathbf{x}$. We distinguish between LOS and NLOS conditions. In the LOS conditions, all the links are considered in LOS. In the NLOS conditions, the probability for each link of being in NLOS is 0.3. We simulate a 5G NR downlink signalling for DL-TDOA. All gNBs are synchronized to transmit the positioning reference signal (PRS) at the same time. The UE estimates the time-of-arrival (TOA) of the PRS signals from different gNBs, and RSTD values are computed for all the pairs of neighboring gNB and reference gNB. We consider PRS signals with sub-carrier spacing of 30 KHz at 3 GHz. The transmit power is 44 dBm, the transmitting and receiving antenna gains are set to 5 dB and 3 dB, respectively. The noise figure is set to 6 dB. For each gNB-UE link, the link-level fading channel is modelled according to the standard NR tapped delay line (TDL) as defined in [66], with path loss calculated according to the urban macro (UMA) scenario. Specifically, LOS channels are modelled through the TDL-D and TDL-E, while the NLOS channels are modelled through the TDL-A, TDL-B, and TDL-C.

2) *Processing*: The TOA estimation is obtained through a peak search after cross-correlation between the transmitted and received waveforms [66]. Then, the RSTD is obtained with respect to the reference gNB. MVEs are obtained from SIs

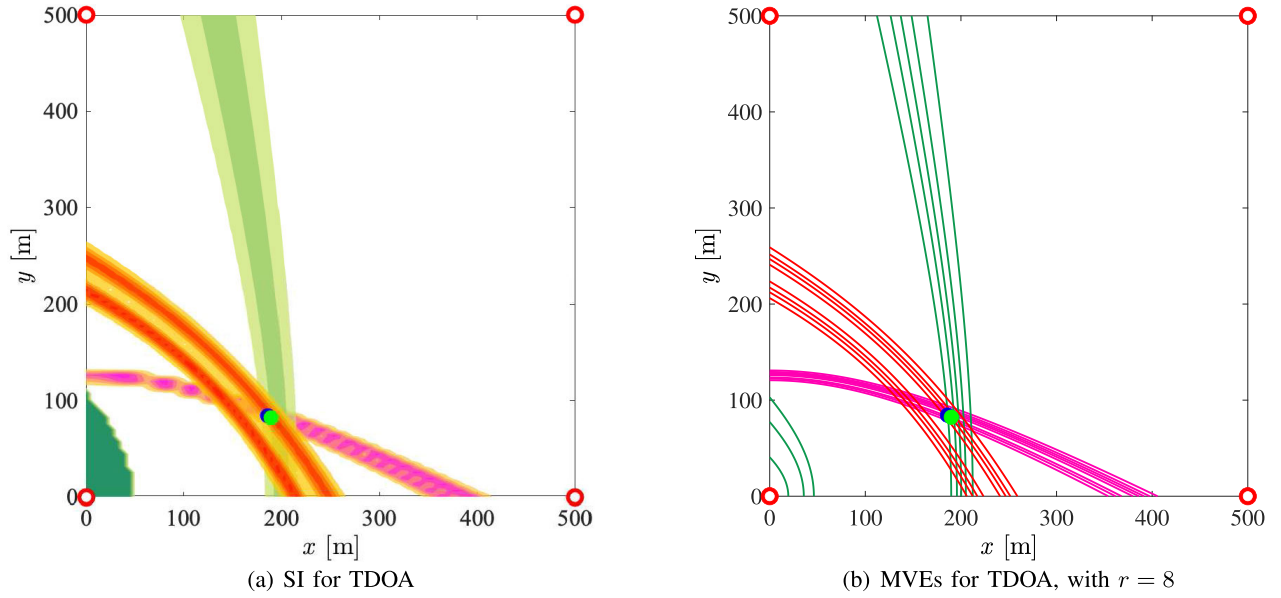


Fig. 7. Graphical representation of SI (a) and MVEs (b) for a single receiver position (blue dot). SI is obtained using GMM, whereas MVEs are obtained using GMM and BC. The gNBs are indicated with red circles and the estimated position is indicated by a green dot.

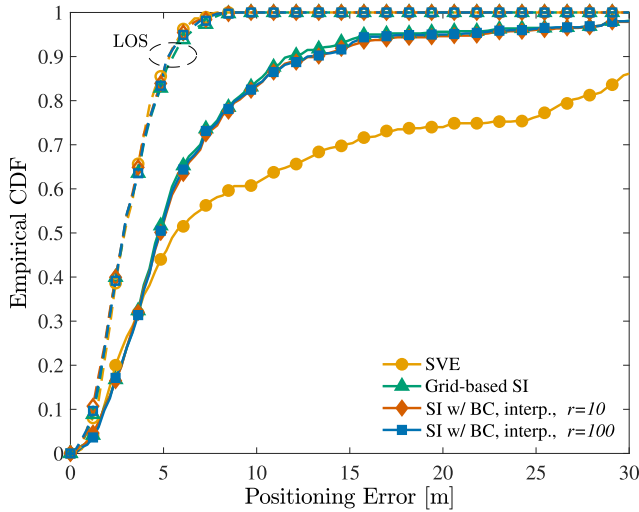


Fig. 8. Empirical CDF of the localization error varying the localization algorithm (SVE, grid-based SI, and SI with BC and statistical interpolation).

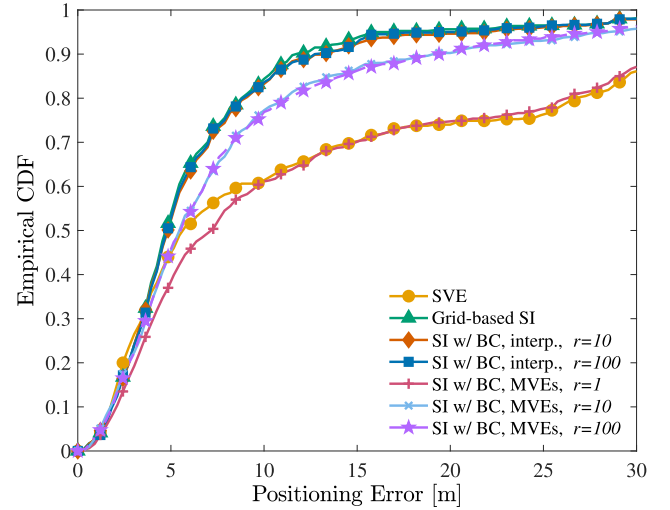


Fig. 9. Empirical CDF of the localization error varying the localization algorithm (SVE, grid-based SI, SI with BC and statistical interpolation, and SI with BC used as MVEs).

as described in Sec. III and then used to estimated positions solving (7) with $l = 2$, with $\Theta_{BC,i} = \{\hat{\theta}_{i,j}\}_{j=1}^r$ being the MVEs for each pair i of gNBs, varying r . The statistical interpolation is evaluated by reconstructing the SI starting from the MVEs $\Theta_{BC,i}$ as described in Sec. III-B.2 using $h = 2r$ in (8). Specifically, in this case study, SI is learned by using a GMM model. The SI model is trained with 200 values of $\mathbf{z}_i = \{y_i, \theta_i\}$ for the three gNB with θ_i being the ground truth RSTD value. The training set is updated at each iteration for each position estimate. However, simulations using a single training set for all iterations did not show significant deviations in performance results. The SVE-based estimation is obtained by solving (2) with $l = 2$ and $\hat{\theta}_i = y_i$ with $i = 1, 2, 3$ being the RSTD measurements. Note that the SVE-based localization is equivalent to the TDOA-based localization as each pair of gNBs yields a single-valued RSTD estimate. Using norm

minimization in (7) with $l = 2$ corresponds to a TDOA-based least squares approach. The grid-based SI is obtained using a number of uniformly equidistant grid points equal to $G = 100^2$ unless specified otherwise.

Figs. 7a and 7b show graphical representations of the SI and the MVEs, respectively, for a given UE position. The green contour in Figure 7a and the green lines in Figure 7b correspond to the top left gNB. It can be seen that the SI captures the multimodality inherent in the TDOA measurements due to the NLOS conditions. Notably, the proposed MVEs are also capable of capturing this multimodality.

Fig. 8 and Fig. 9 show the performance in terms of localization accuracy through the empirical CDF of the localization error, which is given by the Euclidean distance between the true and estimated position. In these figures, we set norm minimization with $l = 1$ for both SVE-based localization

and SI-based localization with MVEs. This choice was made because $l = 1$ and $l = 2$ provide similar performance for SVE-based localization, while $l = 1$ provides better performance for SI-based localization with MVEs. Fig. 8 shows the empirical CDF of the localization error by comparing the SVE, SI and SI with BC and statistical interpolation, the latter with $r = 10$ and $r = 100$. Dashed lines refer to the case with only LOS conditions. The results show that in the LOS conditions, the performance are very similar with SVE, grid-based SI, as well as SI with BC and statistical interpolation. For instance, all the algorithms achieve a 90th percentile localization error of 5.5 m. Differently, in the presence of NLOS conditions with probability 0.3, the performance of SVE-based localization greatly degrades, achieving a 90th percentile localization error of 35 m. In such a case, the usage of SI can mitigate the effect of NLOS conditions and all the SI algorithms achieve a 90th percentile localization error below 14 m. In particular, the grid-based SI and the use of BC with statistical interpolation reach very similar performance. This demonstrates that the information about the measurement distribution can be condensed in a few samples of SI, i.e., $r = 10$.

Fig. 9 compares the two BC methods, i.e., BC and statistical interpolation, and BC used as MVEs (without statistical interpolation). We here recall that the latter enables the use of more efficient localization algorithms, such as norm minimization. The results indicate that utilizing $r = 10$ MVEs yields a 90th percentile localization error of 22 m, with a significant improvement with respect to the SVE method.

Fig. 10 illustrates the probability for the localization error to be lower than 10 m, i.e., the percentage of simulations where the localization error is below this threshold. The figure also presents the computational time for various localization algorithms in both LOS and NLOS conditions. For SI-based localization, a grid-based algorithm is employed with the number of grid points G determining the spatial resolution (higher G values correspond to higher spatial resolution). The results indicate that SVE-based localization and SI-based localization using BC have similar computational times, but SI-based localization with BC can have superior performance in NLOS conditions. The performance of the grid-based SI localization is contingent on the spatial resolution. Increasing the spatial resolution enhances performance but also raises computational costs. In the NLOS conditions, achieving performance comparable to SI with BC (with a 77% probability of the error being below 10 meters) requires 4.3 times larger computational time. If a higher computational time is acceptable (15 times larger with respect to the SVE-based localization), the grid-based SI can even outperform SI-based localization with BC, achieving an 84% probability of maintaining localization error below 10 m.

B. Case Study 2: DOA-Based Localization With Beam Sweeping

The second case study focuses on localization using beam sweeping. This technology can be used for both active and passive users in various applications, such as wireless communications, radar systems, and optical sensing. By sweeping the beam over a specific area, it is possible to detect and locate

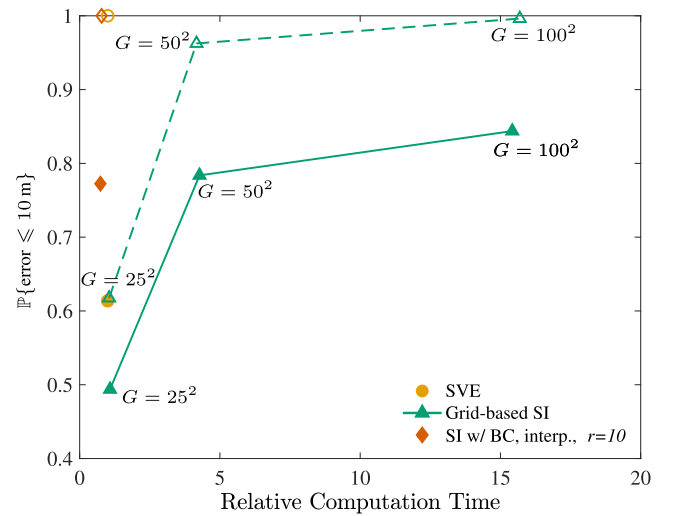


Fig. 10. Empirical probability for the localization error below 10 m in the LOS (dashed) and NLOS (solid) conditions, shown as a function of the computation time. The computation time is normalized to that of SVE-based localization. The grid-based SI localization is shown for different values of grid points G .

objects or signals within that region. In wireless communications, beam sweeping can be used to improve signal strength and coverage by directing the beam towards a specific location or receiver [48], [49], [50], [51]. The beamscan algorithm estimates the DOA by scanning the array beam over a region of interest. The algorithm computes the output power for each beamscan angle over a finite number of possible angles and identifies the maxima as the DOA estimate.

1) *Simulation Settings:* We consider a square scenario with a side length of 40 m, with four receivers positioned at the corners. The transmitter is located at a random and unknown position, denoted as $\mathbf{x}$, uniformly distributed within the square. The receivers utilize a Uniform Linear Array (ULA) comprising ten isotropic antenna elements, operating at a carrier frequency of 1 GHz and with ideal non-overlapping beams. The transmitter is a narrowband source producing linear Frequency Modulated (FM) pulses with 100 KHz bandwidth. The transmit power is 30 dBm and an additive white Gaussian noise channel with noise power of 10 dBm is considered. To simulate NLOS conditions, the incident angle is deviated of a bias uniformly distributed between 10 and 20 degrees with probability 0.5. Each measurement y_i represents the angle estimate for the i th receiver, a discrete variable with a varying number of possible values determined by the size of the antenna array and beam width. In this simulation, we adopt an angular resolution of 12 degrees, resulting in 8 possible values for each angle estimate.

2) *Processing:* SI-based estimation of the transmitter position is obtained through the proposed maximum entropy method and with a training dataset of 200 positions. In this case study, y_i with $i = 1, 2, 3, 4$ are the discrete angle measurements from the four receivers. Then, GMM is not directly applicable since a Gaussian modelling cannot represent the variable y_i that takes only 8 different values. For each receiver, SI training relies on 200 realizations of $\mathbf{z}_i = \{y_i, \theta_i\}$, with θ_i being the ground truth angle values. The sample space for the

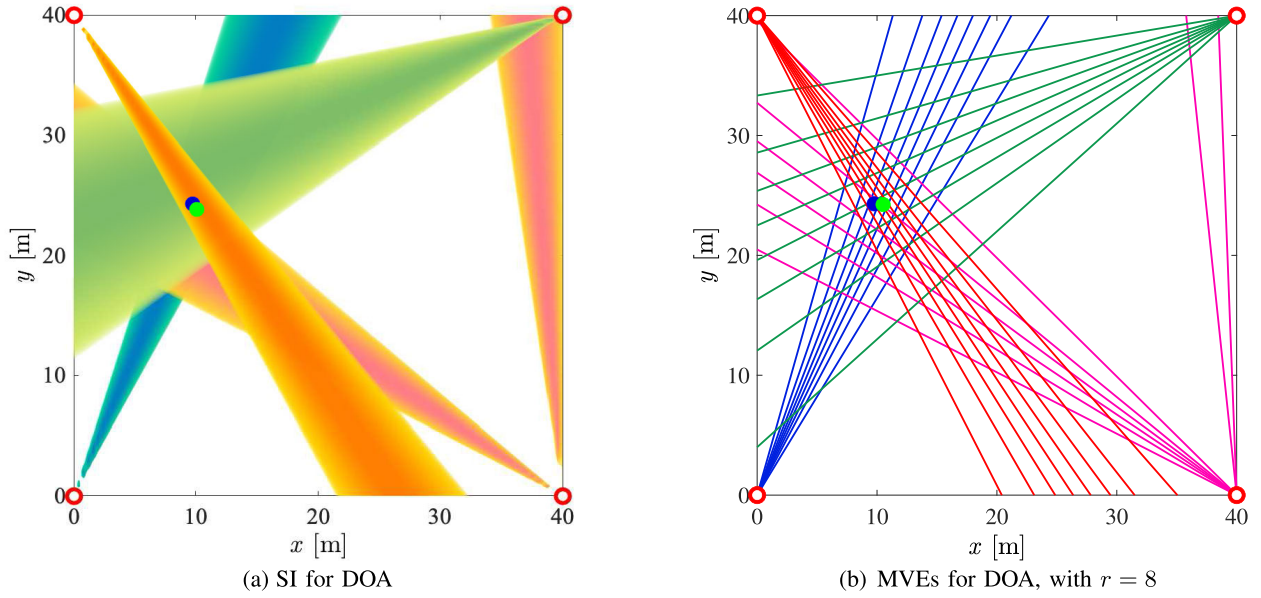


Fig. 11. Graphical representation of SI (a) and MVEs (b) for a single transmitter position (blue dot). SI is obtained using MEM, MVEs are obtained using MEM and BC. The receivers are indicated with red circles and the transmitter estimated position is indicated by a green dot.

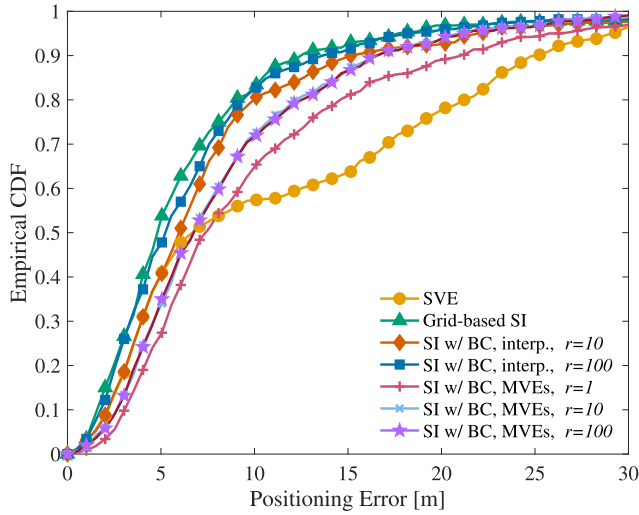


Fig. 12. Empirical CDF of the localization error varying the localization algorithm. The SVE-based localization uses norm minimization with $l = 2$, and SI-based localization with MVEs uses norm minimization with $l = 1$.

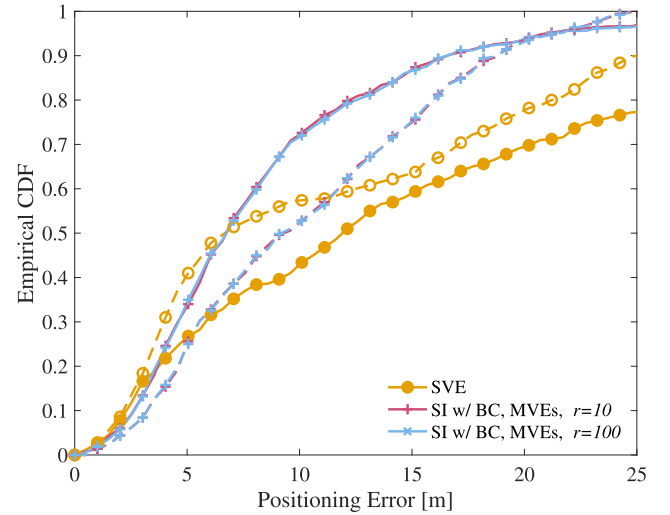


Fig. 13. Empirical CDF of the localization error varying the localization algorithm. The SVE-based localization and MVE-based localization both use norm minimization with $l = 1$ (solid lines) and $l = 2$ (dashed lines).

numerical integral in (13) is a uniform grid of 100 angle values between -180 and 180 degrees. The SI-based localization with MVEs is accomplished by solving (7) with $l = 1$, where $\Theta_{BC} = \{\hat{\theta}_{i,j}\}_{j=1}^r$ are the output values of BC. Then, the statistical interpolation is evaluated by reconstructing the SI starting from the same values Θ_{BC} . SVE-based localization is accomplished by solving (2) with $l = 2$. This choice was made to ensure a fair comparison. As demonstrated later in Fig. 13, comparing the results for $l = 1$ and $l = 2$, we observe that $l = 1$ performs better for SI-based localization with MVEs due to its robustness against outliers, whereas the SVE-based localization yields better results with $l = 2$.

Figs. 11a and 11b show graphical representations of the SI and the MVEs, respectively, for a given transmitter position. The red contour in Figure 11a and the red lines in Figure 11b correspond to top left receiver. It can be seen that the SI

captures the uncertainty of the DOA measurements due to the beamwidth and the power distribution among the different beamscan angles. Notably, the proposed MVEs are also capable of capturing this uncertainty.

Figs. 12 and 13 show the localization performance in terms of empirical CDF of the localization error, i.e., the Euclidean distance between the true and estimated UE position. Fig. 12 shows the empirical CDF of the localization error for different implementations of SI-based localization. The classical SVE-based localization is used as benchmark. The figure shows that the grid-based SI provides the best performance. In addition, SI-based localization with BC and statistical interpolation provides similar performance, especially for high values of r . For instance, both grid-based SI and SI with BC and statistical interpolation ($r = 100$) achieve a 90th percentile localization error of 13 m. The percentile increases to 15 m

when $r = 10$ with statistical interpolation and to 16 m without interpolation (either $r = 10$ or $r = 100$). Yet, the 90-percentile for classical SVE-based localization is 26 m. In addition, SI-based localization with MVEs results in a significantly improved performance compared to SVE-based localization. The figure also shows that such performance improvement can be further increased by using interpolation methods, at the expenses of a higher computational complexity.

Fig. 13 compares the SVE-based localization using (2) and the SI-based localization with MVEs using (7) both with $l = 1$ and $l = 2$. The results show that the norm minimization with $l = 1$ outperforms the $l = 2$ case. Furthermore, the impact of l is much larger for MVE-based localization. The use of norm minimization with $l = 1$ is especially suitable for SI-based localization with MVEs as such optimization is robust to the presence of outliers that are more frequent using multiple values. For example, the 90-percentile of the localization error is 16 m for SI-based localization with MVEs with $l = 1$ and increases to 19 m for $l = 2$. For SVE-based localization, there is a crossing between the two curves, the 90-percentile of the localization error is 25 m for SVE-based localization with $l = 2$ and increases to 36 m for $l = 1$.

VI. FINAL REMARK

The paper presented a framework to enhance both efficiency and accuracy of the localization process by leveraging soft-information (SI). Through the incorporation of alternative machine learning techniques and belief condensation, the presented methods offer two key benefits: (i) the reduction of communication and computing loads through belief condensation, and (ii) the facilitation of handling discrete or categorical measurements through the maximum entropy method. Results from two case studies, which compared algorithms of varying complexity, demonstrated that SI-based localization can achieve significant improvements in localization accuracy with reduced computational and communication demands. By integrating belief condensation and maximum entropy methods, the presented techniques enhance the versatility of SI-based localization, providing a promising solution for applications that require a reduced computational overhead together with an increased adaptability to diverse sensing measurements.

REFERENCES

- [1] M. Z. Win et al., "Network localization and navigation via cooperation," *IEEE Commun. Mag.*, vol. 49, no. 5, pp. 56–62, May 2011.
- [2] M. Z. Win, Y. Shen, and W. Dai, "A theoretical foundation of network localization and navigation," *Proc. IEEE*, vol. 106, no. 7, pp. 1136–1165, Jul. 2018.
- [3] K. Pahlavan, X. Li, and J. P. Makela, "Indoor geolocation science and technology," *IEEE Commun. Mag.*, vol. 40, no. 2, pp. 112–118, Feb. 2002.
- [4] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.
- [5] A. H. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: Challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005.
- [6] *Technical Specification Group Services and System Aspects; Service Requirements for the 5G System*, Standard TS 22.261 V19.6.0, 3rd Gener. Partnership Project, Mar. 2024.
- [7] *Technical Specification Group Services System Aspects; Service Requirements for Cyber-physical Control Applications in Vertical Domains*, Standard TS 22.104 V19.1.0, 3rd Gener. Partnership Project, Sep. 2023.
- [8] *Technical Specification Group Radio Access Network; Study on Scenarios and Requirements of In-coverage, Partial Coverage, and Out-of-coverage NR Positioning Use Cases*, Standard TR 38.845 V17.0.0, 3rd Gener. Partnership Project, Sep. 2021.
- [9] A. Conti et al., "Location awareness in beyond 5G networks," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 22–27, Nov. 2021.
- [10] 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.
- [11] 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.
- [12] 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, Aug. 2016.
- [13] 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.
- [14] 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.
- [15] K. Witrals et al., "High-accuracy localization for assisted living," *IEEE Signal Process. Mag.*, vol. 33, no. 2, pp. 59–70, Mar. 2016.
- [16] M. Chiani, A. Giorgetti, and E. Paolini, "Sensor radar for object tracking," *Proc. IEEE*, vol. 106, no. 6, pp. 1022–1041, Jun. 2018.
- [17] S. Bartoletti, A. Conti, and M. Z. Win, "Device-free counting via wideband signals," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 5, pp. 1163–1174, May 2017.
- [18] X. Ying, S. Roy, and R. Poovendran, "Pricing mechanisms for crowd-sensed spatial-statistics-based radio mapping," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 2, pp. 242–254, Jun. 2017.
- [19] R. Estrada, R. Mizouni, H. Otrok, A. Ouali, and J. Bentahar, "A crowd-sensing framework for allocation of time-constrained and location-based tasks," *IEEE Trans. Services Comput.*, vol. 13, no. 5, pp. 769–785, Sep. 2020.
- [20] D. Zhang, S. Zhao, L. T. Yang, M. Chen, Y. Wang, and H. Liu, "NextMe: Localization using cellular traces in Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 11, no. 2, pp. 302–312, Apr. 2015.
- [21] 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.
- [22] 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.
- [23] *Evolved Universal Terrestrial Radio Access (E-UTRA); Study on Minimization of Drive-tests in Next Generation Networks*, Standard TR 36.805 V9.0.0, 3rd Gener. Partnership Project, Jan. 2010.
- [24] 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.
- [25] S. Bartoletti et al., "Location-based analytics in 5G and beyond," *IEEE Commun. Mag.*, vol. 59, no. 7, pp. 38–43, Jul. 2021.
- [26] Y. Han, Y. Shen, X.-P. Zhang, M. Z. Win, and H. Meng, "Performance limits and geometric properties of array localization," *IEEE Trans. Inf. Theory*, vol. 62, no. 2, pp. 1054–1075, Feb. 2016.
- [27] Y. Li, S. Mazuelas, and Y. Shen, "A variational learning approach for concurrent distance estimation and environmental identification," *IEEE Trans. Wireless Commun.*, vol. 22, no. 9, pp. 6252–6266, Sep. 2023.
- [28] Z. Wang, Z. Liu, Y. Shen, A. Conti, and M. Z. Win, "Location awareness in beyond 5G networks via reconfigurable intelligent surfaces," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2111–2125, Jul. 2022.
- [29] S. Bartoletti et al., "Positioning and sensing for vehicular safety applications in 5G and beyond," *IEEE Commun. Mag.*, vol. 59, no. 11, pp. 15–21, Nov. 2021.

- [30] S. Bartoletti, N. Decarli, and B. M. Masini, "Sidelink 5G-V2X for integrated sensing and communication: The impact of resource allocation," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, May 2022, pp. 79–84.
- [31] G. Kwon, Z. Liu, A. Conti, H. Park, and M. Z. Win, "Integrated localization and communication for efficient millimeter wave networks," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 12, pp. 3925–3941, Dec. 2023.
- [32] L. Pucci, E. Paolini, and A. Giorgetti, "System-level analysis of joint sensing and communication based on 5G new radio," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 7, pp. 2043–2055, Jul. 2022.
- [33] S. Bartoletti, Z. Liu, M. Z. Win, and A. Conti, "Device-free localization of multiple targets in cluttered environments," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 58, no. 5, pp. 3906–3923, Oct. 2022.
- [34] S. Bartoletti, W. Dai, A. Conti, and M. Z. Win, "A mathematical model for wideband ranging," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 2, pp. 216–228, Mar. 2015.
- [35] J. Park et al., "Communication-efficient and distributed learning over wireless networks: Principles and applications," *Proc. IEEE*, vol. 109, no. 5, pp. 796–819, May 2021.
- [36] F. Gini, F. Lombardini, and P. K. Varshney, "On distributed signal detection with multiple local free parameters," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 35, no. 4, pp. 1457–1466, Apr. 1999.
- [37] G. Torsoli, M. Z. Win, and A. Conti, "Blockage intelligence in complex environments for beyond 5G localization," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 6, pp. 1688–1701, Jun. 2023.
- [38] M. Chen et al., "Distributed learning in wireless networks: Recent progress and future challenges," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 12, pp. 3579–3605, Dec. 2021.
- [39] Z. Liu, A. Conti, S. K. Mitter, and M. Z. Win, "Communication-efficient distributed learning over networks—Part I: Sufficient conditions for accuracy," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1081–1101, Apr. 2023.
- [40] Z. Liu, A. Conti, S. K. Mitter, and M. Z. Win, "Communication-efficient distributed learning over networks—Part II: Necessary conditions for accuracy," *IEEE J. Sel. Areas Commun.*, vol. 41, no. 4, pp. 1102–1119, Apr. 2023.
- [41] 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.
- [42] H. Wymeersch, S. Marano, W. M. Gifford, and M. Z. Win, "A machine learning approach to ranging error mitigation for UWB localization," *IEEE Trans. Commun.*, vol. 60, no. 6, pp. 1719–1728, Jun. 2012.
- [43] Y. Huang, S. Mazuelas, F. Ge, and Y. Shen, "Indoor localization system with NLOS mitigation based on self-training," *IEEE Trans. Mobile Comput.*, vol. 22, no. 7, pp. 3952–3966, Jul. 2023.
- [44] 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, Sep. 2019.
- [45] S. Mazuelas, A. Conti, J. C. Allen, and M. Z. Win, "Soft range information for network localization," *IEEE Trans. Signal Process.*, vol. 66, no. 12, pp. 3155–3168, Jun. 2018.
- [46] F. Morselli, S. M. Razavi, M. Z. Win, and A. Conti, "Soft information based localization for 5G networks and beyond," *IEEE Trans. Wireless Commun.*, vol. 22, no. 12, pp. 9923–9938, Dec. 2023.
- [47] S. Bartoletti et al., "Positioning integrity via uncertainty quantification," *IEEE Trans. Veh. Technol.*, vol. 74, no. 1, pp. 1224–1237, Jan. 2025.
- [48] G. Fokin, "Bearing measurement with beam sweeping for positioning in 5G networks," in *Proc. IEEE Microw. Theory Techn. Wireless Commun. (MTTW)*, Riga, Latvia, Oct. 2021, pp. 64–67.
- [49] Y. Zhou, G. Liu, J. Li, Y. Li, S. Ye, and L. Li, "A high-efficiency beam sweeping algorithm for DOA estimation in the hybrid analog-digital structure," *IEEE Wireless Commun. Lett.*, vol. 10, no. 10, pp. 2323–2327, Oct. 2021.
- [50] H. Ding and K. G. Shin, "Accurate angular inference for 802.11ad devices using beam-specific measurements," *IEEE Trans. Mobile Comput.*, vol. 21, no. 3, pp. 822–834, Mar. 2022.
- [51] T. Izydorczyk, G. Berardinelli, F. M. L. Tavares, M. Bucur, and P. Mogensen, "On the potential of uplink beamforming in vehicular networks based on experimental measurements," in *Proc. IEEE 90th Veh. Technol. Conf. (VTC-Fall)*, Sep. 2019, pp. 1–5.
- [52] C. Qin, J. A. Zhang, X. Huang, K. Wu, and Y. J. Guo, "Fast angle-of-arrival estimation via virtual subarrays in analog antenna array," *IEEE Trans. Wireless Commun.*, vol. 19, no. 10, pp. 6425–6439, Oct. 2020.
- [53] J. A. Zhang, W. Ni, P. Cheng, and Y. Lu, "Angle-of-arrival estimation using different phase shifts across subarrays in localized hybrid arrays," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2205–2208, Nov. 2016.
- [54] K. Wu, W. Ni, T. Su, R. P. Liu, and Y. J. Guo, "Robust unambiguous estimation of angle-of-arrival in hybrid array with localized analog sub-arrays," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2987–3002, May 2018.
- [55] M. Gavish and A. J. Weiss, "Performance analysis of bearing-only target location algorithms," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 817–828, Jul. 1992.
- [56] S. Mazuelas, Y. Shen, and M. Z. Win, "Belief condensation filtering," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4403–4415, Sep. 2013.
- [57] S. Mazuelas, Y. Shen, and M. Z. Win, "Belief condensation filter for navigation in harsh environments," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Kyoto, Japan, Jun. 2011, pp. 1–6.
- [58] S. Mazuelas, Y. Shen, and M. Z. Win, "Framework of belief condensation filtering and deterministic discrete filters," in *Proc. IEEE Int. Conf. Ultra-Wideband*, Sep. 2012, pp. 1–5.
- [59] E. T. Jaynes, "Information theory and statistical mechanics," *Phys. Rev.*, vol. 106, no. 4, pp. 620–630, 1957.
- [60] E. T. Jaynes, *Probability Theory: The Logic of Science*. Cambridge, U.K.: Cambridge Univ. Press, 2003.
- [61] J. Shore and R. Johnson, "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Trans. Inf. Theory*, vol. IT-26, no. 1, pp. 26–37, Jan. 1980.
- [62] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*, 2nd ed., Cambridge, MA, USA: MIT Press, 2018.
- [63] M. Dudík, S. J. Phillips, and R. E. Schapire, "Performance guarantees for regularized maximum entropy density estimation," in *Proc. Conf. Comput. Learn. Theory*, Berlin, Germany, Jul. 2004, pp. 472–486.
- [64] S. J. Phillips, R. P. Anderson, and R. E. Schapire, "Maximum entropy modeling of species geographic distributions," *Ecological Model.*, vol. 190, nos. 3–4, pp. 231–259, Jan. 2006.
- [65] S. Mazuelas, Y. Shen, and A. Pérez, "Generalized maximum entropy for supervised classification," *IEEE Trans. Inf. Theory*, vol. 68, no. 4, pp. 2530–2550, Apr. 2022.
- [66] *Technical Specification Group Radio Access Network; Study on Channel Model for Frequencies From 0.5 To 100 GHz*, Standard TR 38.901 V17.0.0, 3GPP, Mar. 2022.
- [67] A. Conti, G. Torsoli, C. A. Gómez-Vega, A. Vaccari, G. Mazzini, and M. Z. Win, "3GPP-compliant datasets for xG location-aware networks," *IEEE Open J. Veh. Technol.*, vol. 5, pp. 473–484, 2024.
- [68] A. Conti, G. Torsoli, C. A. Gómez-Vega, A. Vaccari, and M. Z. Win, Dec. 2023, "xG-Loc: 3GPP-compliant datasets for xG location-aware networks," IEEE Dataport, doi: [10.21227/rper-vc03](https://doi.org/10.21227/rper-vc03).
- [69] Y. Chen, H. C. So, and E. E. Kuruoglu, "Variance analysis of unbiased least ℓ_p -norm estimator in non-Gaussian noise," *Signal Process.*, vol. 122, pp. 190–203, May 2016.
- [70] S. A. Flores, "Robustness of ℓ_1 -norm estimation: From folklore to fact," *IEEE Signal Process. Lett.*, vol. 25, no. 11, pp. 1640–1644, Nov. 2018.
- [71] H. L. Van Trees, *Detection, Estimation, and Modulation Theory*, 1st ed., New York, NY, USA: Wiley, 1968.
- [72] S. Dasgupta, "Learning mixtures of Gaussians," in *Proc. 40th Annu. Symp. Found. Comput. Sci.*, 1999, pp. 634–644.
- [73] A. Bakshi, I. Diakonikolas, H. Jia, D. M. Kane, P. K. Kothari, and S. S. Vempala, "Robustly learning mixtures of arbitrary Gaussians," in *Proc. 54th Annu. ACM SIGACT Symp. Theory Comput.*, Rome, Italy, Jun. 2022, pp. 1234–1247.
- [74] C. O'Meadhra, W. Tabib, and N. Michael, "Variable resolution occupancy mapping using Gaussian mixture models," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 2015–2022, Apr. 2019.
- [75] R. J. Kozick and B. M. Sadler, "Maximum-likelihood array processing in non-Gaussian noise with Gaussian mixtures," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3520–3535, Dec. 2000.
- [76] J. L. Devore, *Probability and Statistics for Engineering and the Sciences*, 8th ed., Boston, CA, USA: Brooks/Cole, 2011.
- [77] L. Yu, T. Yang, and A. B. Chan, "Density-preserving hierarchical EM algorithm: Simplifying Gaussian mixture models for approximate inference," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 6, pp. 1323–1337, Jun. 2019.
- [78] E. Coviello, A. B. Chan, and G. Lanckriet, "Clustering hidden Markov models with variational HEM," *ACM J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 697–747, Jan. 2014.

- [79] M. A. Brubaker, A. Geiger, and R. Urtasun, "Map-based probabilistic visual self-localization," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 4, pp. 652–665, Apr. 2016.
- [80] K. Zhang and J. T. Kwok, "Simplifying mixture models through function approximation," *IEEE Trans. Neural Netw.*, vol. 21, no. 4, pp. 644–658, Apr. 2010.
- [81] I. Csiszar, "I-divergence geometry of probability distributions and minimization problems," *Ann. Probab.*, vol. 3, pp. 146–158, Feb. 1975.
- [82] T. Sutter, D. Sutter, P. M. Esfahani, and J. Lygeros, "Generalized maximum entropy estimation," *J. Mach. Learn. Res.*, vol. 20, no. 138, pp. 1–29, Sep. 2019.
- [83] T. Saad and G. Ruai, "PyMaxEnt: A Python software for maximum entropy moment reconstruction," *SoftwareX*, vol. 10, Jul. 2019, Art. no. 100353.
- [84] Y. Altun and A. Smola, "Unifying divergence minimization and statistical inference via convex duality," in *Proc. Annu. Conf. Learn. Theory*, 2006, pp. 139–153.
- [85] D. J. MacKay, *Information Theory, Inference, and Learning Algorithms*. New York, NY, USA: Cambridge Univ. Press, 2003.



Stefania Bartoletti (Member, IEEE) received the Laurea degree (summa cum laude) in electronics and telecommunications engineering and the Ph.D. degree in information engineering from the University of Ferrara, Italy, in 2011 and 2015, respectively.

Dr. Bartoletti is an Associate Professor at University of Rome "Tor Vergata," Italy. From 2019 to 2022, she was a Researcher at the Institute of Electronics, Computer and Telecommunication Engineering, National Research Council of Italy (IEIT-CNR). From 2016 to 2019, she was a Marie

Sklodowska-Curie Global Fellow within the Horizon 2020 European Framework at Massachusetts Institute of Technology, Cambridge, MA, USA, and the University of Ferrara. Her research interests include theory and experimentation of wireless networks for sensing and localization. She was awarded a Starting Grant in 2023 from European Research Council. She served as the Chair of the Technical Program Committee (TPC) for several IEEE ICC and Globecom Workshops. She is an Editor of IEEE COMMUNICATIONS LETTERS.



Santiago Mazuelas (Senior Member, IEEE) received the Ph.D. degree in mathematics and the Ph.D. degree in telecommunications engineering from the University of Valladolid, Spain, in 2009 and 2011, respectively. From 2009 to 2014, he was a Post-Doctoral Fellow and an Associate Researcher at the Wireless Information and Network Sciences Laboratory, Massachusetts Institute of Technology (MIT).

Dr. Mazuelas is an Ikerbasque Associate Professor at the Basque Center for Applied Mathematics (BCAM). Prior to joining BCAM, he was a Staff Engineer at Qualcomm Research and Development from 2014 to 2017. He has received the Young Scientist Prize from the Union Radio-Scientifique Internationale (URSI) Symposium in 2007 and the Early Achievement Award from the IEEE ComSoc in 2018. His papers have received the IEEE Communications Society Fred W. Ellersick Prize in 2012, the SEIO-FBBVA Best Applied Contribution in the Statistics Field in 2022, and the Best Paper Award from the IEEE ICC, the IEEE ICUWB, and the IEEE Globecom. He has served as the Technical Program Vice-Chair for the 2021 IEEE Globecom as well as the Symposium Co-Chair for the 2014 IEEE Globecom, the 2015 IEEE ICC, and the 2020 IEEE ICC. He was an Area Editor of IEEE COMMUNICATION LETTERS from 2017 to 2022, and he is currently an Associate Editor for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and IEEE TRANSACTIONS ON MOBILE COMPUTING.



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

In Summer 2001, he was with the Wireless Systems Research Department at AT&T Research Laboratories. Since 2003, he has been a frequent visitor with the Wireless Information and Network Sciences Laboratory, Massachusetts Institute of Technology, where he presently holds the Research

Affiliate appointment. His research involves the theory and experimentation of wireless communication and localization systems. His current 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 and 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 IET and a member of Sigma Xi. He has been selected as an IEEE Distinguished Lecturer.



Moe Z. Win (Fellow, IEEE) is the Robert R. Taylor 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 the 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 for 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. He has served on the SIAM Diversity Advisory Committee.

Dr. Win is elected as a 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 MIT Frank E. Perkins Award (2024), the MIT Everett Moore Baker Award (2022), the IEEE Vehicular Technology Society James Evans Avant Garde Award (2022), 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).