CHALMERS
UNIVERSITY OF TECHNOLOGY

RiLoCo: An ISAC-oriented AI Solution to Build RIS-empowered Networks

Citation for the original published paper (version of record):
Encinas-Lago, G., Sciancalepore, V., Wymeersch, H. et al (2025). RiLoCo: An ISAC-oriented AI Solution to Build RIS-empowered Networks. IEEE Transactions on Wireless Communications, 24(10): 8221-8235. http://dx.doi.org/10.1109/TWC.2025.3564832

(article starts on next page)

RiLoCo: An ISAC-oriented AI Solution to Build RIS-empowered Networks

Guillermo Encinas-Lago, *Student Member, IEEE,* Vincenzo Sciancalepore, *Senior Member, IEEE,*
Henk Wymeersch, *Fellow, IEEE,* Marco Di Renzo, *Fellow, IEEE,*
Xavier Costa-Pérez, *Senior Member, IEEE*

***Abstract*—The advance towards 6G networks comes with the promise of unprecedented performance in sensing and communication capabilities. The feat of achieving those, while satisfying the ever-growing demands placed on wireless networks, promises revolutionary advancements in sensing and communication technologies. As 6G aims to cater to the growing demands of wireless network users, the implementation of intelligent and efficient solutions becomes essential. In particular, reconfigurable intelligent surfaces (RISs), also known as Smart Surfaces, are envisioned as a transformative technology for future 6G networks.**

The performance of RISs when used to augment existing devices is nevertheless largely affected by their precise location. Suboptimal deployments are also costly to correct, negating their low-cost benefits. This paper investigates the topic of optimal RISs diffusion, taking into account the improvement they provide both for the sensing and communication capabilities of the infrastructure while working with other antennas and sensors. We develop a combined metric that takes into account the properties and location of the individual devices to compute the performance of the entire infrastructure. We then use it as a foundation to build a reinforcement learning architecture that solves the RIS deployment problem. Since our metric measures the surface where given localization thresholds are achieved and the communication coverage of the area of interest, the novel framework we provide is able to seamlessly balance sensing and communication, showing its performance gain against reference solutions, where it achieves simultaneously almost the reference performance for communication and the reference performance for localization.

I. Introduction

Over the past decades, wireless communications and radar sensing have made significant advancements as prominent radio technologies. However, these technologies have developed independently with a very limited collaboration, despite their shared hardware architecture and signal processing principles [1]. Recently, the idea of integrated sensing and communication (ISAC) has emerged, garnering considerable attention in the research community [2]. ISAC aims to combine the functionalities of communication and sensing systems, leveraging scarce spectral resources to create cost-efficient, intelligent, and user-friendly wireless networks with unprecedented possibilities [3]. Nevertheless, establishing a reliable ISAC service poses serious challenges, particularly in unfavorable electromagnetic (EM) propagation environments, operating at millimeter-wave (mmWave) and terahertz (THz) frequency bands, which offer abundant spectral resources [4]. To overcome such limitations, researchers are currently exploring more sophisticated solutions: On the one hand, compensating for the path loss in high-frequency channels can be achieved through the use of massive multiple-input, multiple-output (MIMO) systems [5], which provide high beamforming and spatial gains; on the other hand, RISs have shown potential in mitigating the blockage effect among other problems [6]. Indeed, the groundbreaking RIS technology can programmatically alter the propagation properties of an incoming EM signal, thereby providing the means to control the surrounding propagation environment. Despite such promising techniques being extensively studied in separate fields of communications and sensing, their specific applications to ISAC systems still remain relatively unexplored.

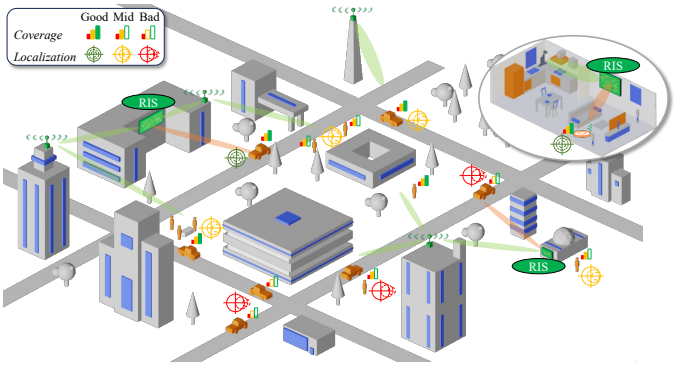

Fig. 1. RIS-assisted ISAC network simultaneously considering both indoor and outdoor scenarios.

RISs, proven to be low-complexity, low-cost, and nearly passive devices, can play a major role in accomplishing the ISAC revolution due to their sustainability and flexibility properties. However, installing and maintaining such innovative devices can become an impelling task. RISs must be properly placed and continuously configured to simultaneously i) steer the communication beam towards the desired receiver and ii) exploit the passive radar capabilities to collect real-time information from connected devices, which send timestamped

Guillermo Encinas-Lago is with NEC Laboratories Europe, 69115 Heidelberg, Germany; Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes, 3 Rue Joliot-Curie, 91192 Gif-sur-Yvette, France; and i2Cat, 08034 Barcelona, Spain.
Vincenzo Sciancalepore is with NEC Laboratories Europe.
Henk Wymeersch is with Chalmers University of Technology, Department of Electrical Engineering, 412 96 Gothenburg, Sweden.
Marco Di Renzo is with Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des Signaux et Systèmes. (marco.di-renzo@universite-paris-saclay.fr), and with King's College London, Centre for Telecommunications Research – Department of Engineering, WC2R 2LS London, United Kingdom (marco.di_renzo@kcl.ac.uk).
Xavier Costa-Pérez is with i2Cat, ICREA, and NEC Laboratories Europe.
Email of the corresponding author: guillermo.encinas@i2cat.net.

pilot signals and infer the target position, as depicted in Fig. 1.

Nonetheless, during the initial roll-out phase, the placement of RISs becomes of paramount importance: this has been recently studied in previous works that tackle the RIS deployment problem [7], [8] from a theoretical perspective suggesting practical frameworks. In those practical frameworks, the performance gains provided by RISs vary greatly with their placement. Additionally, the solutions to the deployment problem (i.e. the locations where RISs are placed) can often be counterintuitive. These two properties (the importance and the counterintuitiveness) of the deployment solutions, call for solid approaches to the issue. While existing works attain to communications capabilities boost, they neglect other important benefits that networks obtain from incorporating RISs. Conversely, in this work, we specifically aim to cover that gap, building a localization metric in seamless cooperation with a communications metric. We rely on such novel key performance indicators (KPIs) while devising a practical RIS installation framework, as detailed in the following section.

Contributions: In this paper, we propose *RiLoCo*, a reinforcement learning (RL)-based method that, unlike existing solutions, maximizes RIS performance both in sensing and communication. In RiLoCo we explicitly consider environmental signal blockages and shadowing and treat RISs and base stations (BSs) as sensors. To build it, we first transform the information from the sensors into consolidated spatial probability density functions (PDFs) of the user location estimations. Finally, computing the Fisher information contained in those PDFs, we introduce meaningful metrics that serve as the heart of RiLoCo, jointly evaluating the localization and communication performance of deployments while allowing heterogeneous combinations of RIS, BSs, and other sensors.

Contributions are listed as follows: C1) a novel formulation of the RIS deployment problem oriented to RL, catering both sensing and communications performance, considering the simultaneous deployment of other devices (as BSs) C2) mathematical models to describe the information obtainable from RISs, predicting the device capabilities in an accurate and tractable manner, C3) a processing method to encode all the available information from sensors into a single Cartesian discrete PDF of position estimations, C4) a localization performance evaluation of the infrastructure based on Fisher information, able to drive the RL algorithm, C5) a joint sensing and communication metric for RIS, BS and sensor deployments and C6) the performance evaluation of RiLoCo, leveraging on the previously listed contributions in a sample scenario.

Notation and conventions: We denote vectorial quantities with boldface, and the modulus of an N-dimensional vector $\mathbf{v}$ as $\|\mathbf{v}\|$. We use Cartesian coordinates when not otherwise noted. A notable exception is angular measurements of the direction of an impinging signal in an antenna, which we indicate using horizontal coordinates. In this context, we denote the *elevation of arrival* by φ_A and define it as the angle between the arrival direction and the reference plane XY. We denote the *azimuth of arrival* by θ_A and define it as the signed angle between the reference direction X and the projection of the arrival direction in the reference plane XY. Consistent with this, for simplicity we use the function $\text{atan2}(y, x)$, also in the vectorial form $\text{atan2}(\mathbf{p}) \equiv \text{atan2}(y, x)$ with $\mathbf{p} = (x, y, z)$, defined as the following:

$$\text{atan2}(y,x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0, \\ \frac{\pi}{2} - \arctan\left(\frac{x}{y}\right) & \text{if } y > 0, \\ -\frac{\pi}{2} - \arctan\left(\frac{x}{y}\right) & \text{if } y < 0, \\ \arctan\left(\frac{y}{x}\right) \pm \pi & \text{if } x < 0, \\ \text{undefined} & \text{if } x = y = 0. \end{cases}$$

We use the Gaussian distribution, denoted as $\mathcal{N}(\mu, \sigma^2)$, where μ is the expected value or mean of the distribution and σ^2 is the variance. We use c for the speed of light in vacuum. We denote the natural logarithm of a quantity q as $\ln(q)$ or simply $\ln q$, and the average of the same quantity as $\overline{q}$. For the expected value of a function $f(Z)$ of a random variable Z we use the notation $\mathbb{E}[f(Z)]$. The i-th eigenvalue of a matrix M is denoted as $ev_i(M)$.

II. Related Works

The topic of ISAC is currently receiving noticeable attention from academic literature. We summarize here the publications that present the closest parallelism and relevance to the deployment of devices capable of simultaneous sensing and communications (and in particular, RISs), in ISAC systems.

A. ISAC performance

The new trade-offs that arise when considering simultaneously sensing and communications aspects in the infrastructure are studied in works as [9] and [10]. These are taken into consideration when the theoretical limits of the capabilities of ISAC networks are studied as in [9], which can directly impact the deployment problem and other parallel questions, such as the power allocation problem studied in [10].

B. Deployment problem in ISAC systems

Very limited literature has addressed the optimal deployment problem considering both communication and sensing. For instance, [11] studied an optimal BS deployment solution that jointly maximizes the overall system throughput and minimizes the Cramér–Rao bound (CRB), thus improving the localization accuracy. We go one step beyond, looking at RIS as the emerging and groundbreaking technology. It has substantially attracted academic and industrial entities due to its versatility and agility to enhance existing network deployments. In the last few years, a plethora of research works have been published to tackle such novel technologies from different perspectives, including physical modeling, design and implementation, and optimization of RIS configuration and control. However, commercial exploitation involves a number of technical challenges due to the wide set of variables to take into account, such as operating frequency, size and orientation, or optimized KPIs [12].

C. RIS usage in ISAC systems

The exploitation of the reconfigurable propagation environment that RISs are able to build in the context of ISAC systems is also well studied in the literature. For example, in [13] the authors describe the topic and the capabilities RISs offer for ISAC, as well as documenting the state of the art on the techniques studying how to optimally incorporate these devices along existing machines. In [14], the authors proposed a new RIS-aided edge caching system by formulating a network cost minimization problem to jointly optimize content placement at cache units, active beamforming at BS, and RIS configuration. This is performed by exploiting an alternating optimization algorithm to jointly tackle the BS beamforming and RIS configuration.

D. RIS deployment problem

In [15], the authors analyzed the coverage of a downlink RIS-assisted network. The RIS placement optimization problem was formulated to maximize the cell coverage by optimizing the RIS orientation and horizontal distance with the BS. In [16], an optimal RIS placement for highly directional mmWave links is presented. The authors highlighted the relationship between transmission beam footprint at the RIS plane and its corresponding size. The authors of [17] focused on the mmWave communications in indoor scenarios. They formulated an RIS placement optimization to maximize the coverage area with a supervised learning approach outperforming the conventional schemes attained by the decision tree algorithm and the randomly deployed RIS. Finally, [18] evaluated the end-to-end signal-to-noise ratio (SNR) expression of the transmitter-RIS-receiver links to acquire important insights about the RIS position on the overall system performance.

Moving towards higher frequencies, such as D-band, in [19] the optimal placement of the RIS is studied with respect to both position and orientation. In particular, the proposed analysis is based on an analytical model that treats the RIS as a continuous surface of finite size that steers the incident beam toward a prescribed direction.

Nevertheless, none of these approaches looked into the optimal RIS installation and configuration problem to pursue both localization and communication performance maximization at the same time, leading us to work within that niche in the present work.

III. System Model

To define the initial problem and develop our proposed solution, we examine a traditional propagation environment where transmitters and receivers can establish connections via both direct links (i.e., line-of-sight (LoS)) and reflected links (i.e., utilizing the conventional properties of the RIS). We consider a confined scenario, where we intend to deploy devices to improve coverage and localization services while taking into account the existing infrastructure. This is depicted in Fig. 2.

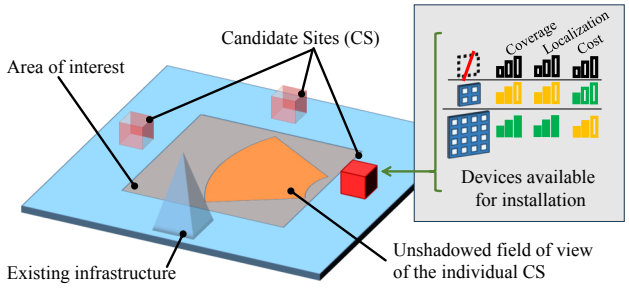

Fig. 2. Schematic view of the elements of the system model.

In this section we explain our *scenario*, detailing the model where we base our work. We explain how we model the devices in the infrastructure able to provide information (*sensing nodes*), and how we encode that information. We consider two particular examples, inspired in RISs: *time of arrival (ToA) sensing* and *angle of arrival (AoA) sensing*. We explain also the way we model *shadowing* in our scenario.

A. Scenario

We study an area where multiple user equipments (UEs) are placed. In this area we want to deliver boosted-performance communication and sensing services by means of a wireless network, as depicted in Fig. 1. To achieve this we consider that we can either install RISs into the environment (in order to enhance the existing wireless infrastructure, e.g., access points or BSs) or by planning a joint BSs and RISs deployment. Our intention is to exploit the controllable (passive) beamforming at the RIS devices that, in turn, become i) effective nearly passive sensors able to accurately detect the position of a target user (please refer to Theorem 1 in Section IV) and ii) nearly passive signal repeaters to significantly enhance the overall communications while ensuring the area receives a sufficient level of coverage at all points. The technical challenge relies on finding the optimal overall network deployment that maximizes both communication and sensing performance.

In our study, we consider devices that are deployed with varying characteristics: this includes differences in the number and arrangement of antennas, the accuracy of their sensing capabilities, and, in the case of BSs, the transmitted power. To better consider the characteristics of the physical environment of our study, we take into account the presence of shadowing between each potential deployment site, namely candidate site (CS) and the area where UEs are distributed. We represent this shadowing information as a collection of blockage areas for each potential CS, and we encode the environment propagation conditions using such blockage areas. This is illustrated in Fig. 2, where the shadowed area of the highlighted CS is colored gray and the unshadowed (usable) area is colored orange. It is also further explained and justified in Sec. III subsection *E. Shadowing and connection models*. Additionally, we can incorporate to our method information about the expected statistical distribution of UE locations in the scenario. In our implementation we consider the worst case for the localization accuracy, a uniform distribution of users, representing the lack of prior information. We exploit the information on shadowing areas to create a model of the propagation environment in

which our solution will operate. This model helps us sir
how each solution will perform in practice. On the other
the statistical distribution of UE locations is utilized
evaluation metrics.[1] By using the statistical distributi
UEs, we can effectively exclude the evaluation of area
are inaccessible or not relevant for our study: this sim
the evaluation process and allows us to focus on the r
where the solutions are viable and applicable with a re
output.

B. Sensing nodes

Fundamentally, the construction of RiLoCo allows f
acceptance of any sensing information (i.e., the output
any sensor) that can be projected into a Cartesian coor
space. This would include, for example, pyroelectric in
presence sensors within a room [20], mmWave radars
channel information fingerprinting [22], or even secondary information sources, such as location information inferred from sound noise levels, energy consumption, etc. For conciseness and clarity, in this work we focus on the deployment of RISs—one of the most relevant technologies for sensing and communications [23]—and BSs. We model and use the information readily obtainable from them as the basis for RiLoCo. We assume that each device (i.e., BSs and RISs) can provide sensing information in the form of either ToA, AoA, or both, as shown in the left and center parts of Fig. 3. On the right part of this figure we also introduce elements properly covered later in the paper, as the combination of the information from all the individual sensors (detailed in Sec. IV) and the extraction of the accuracy achievable from this (detailed in Sec. V. We infer the individual measurement accuracy of each sensor from the intrinsic characteristics (e.g., the number and distribution of elements of an RIS), the operating frequency, the measurement conditions, e.g., the SNR at the sensor doing the measurement, etc. [24] With the accuracy of the sensors, we model the distribution of the measurements expected from each sensor as a Gaussian distribution, where the average μ is the real value of the measured magnitude and σ^2 is the variance of the statistical distribution of the measurements obtained from the sensor. ToA and AoA sensors have different models for the dependence of their accuracy with the SNR. In the following, we describe how ToA and AoA sensing procedures are performed.

C. ToA sensing procedure

We model the BSs placed in the scenario as nodes with ToA sensing capabilities. If we consider a sensor in the location $\mathbf{p}_s = (x_s, y_s, z_s)$ with an internal clock synchronized with the reference time t, and a UE located at $\mathbf{p}_u = (x_u, y_u, z_u)$ with an internal clock t_u presenting an unknown constant deviation from the reference time t so $t - t_u = \Delta_{t_u} \neq 0$, a ToA measurement translates to a spherical cone:

$$(T_A - \Delta_{t_u})^2 \cdot c^2 = (x_u - x_s)^2 + (y_u - y_s)^2 + (z_u - z_s)^2, \quad (1)$$

[1]This can be used, for example, by applying weights to the metric evaluations based on the density of UEs in different spatial regions and collecting corresponding results, as shown in Section VI.

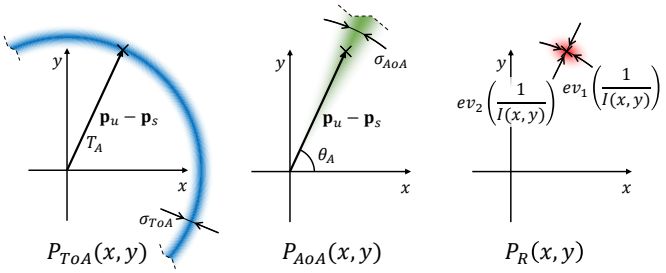

Fig. 3. Time of arrival (ToA) (left) and angle of arrival (AoA) (center) measurements diagrams in 2D, with their combination (right).

where the measurement T_A is the difference between the absolute time of arrival of the sounding signal (in the time reference system) and the UE time of emission of said signal (in the UE own time reference). Eq. (1) presents a simple concept: the right term represents the distance between the UE and the sensor that performs the ToA measurement, squared. The left term represents the time of flight of the measurement signal multiplied by the speed of the signal, c, both squared. This turns into the following simpler form

$$T_A{}^2 \cdot c^2 = (x_u - x_s)^2 + (y_u - y_s)^2, \quad (2)$$

when considering a scenario wherein we are restricted to a known, constant height and where the internal clock deviation of the UE is assumed to be negligible [2].

For the ToA measurement accuracy we take the value of the range accuracy of a secondary radar using a bandwidth B [25]:

$$\sigma_{ToA} = \frac{c}{2 \cdot B\sqrt{2 \cdot \mathrm{SNR}}}. \quad (3)$$

We model RISs as nearly passive devices without sensing or computational capabilities ([12]), and hence they do not offer any direct feedback. Nevertheless, we can obtain implicit information from their operation: the alterations in observable environmental variables—e.g. the channel state information (CSI)—following controlled (or known) changes in the RIS configuration can be exploited to gain knowledge about the environment or, specifically in our case, the position of the UE [6]. We assume the BS is able to discriminate the received signals in order to distinguish those coming from the UE and the ones programmatically reflected by the RIS, thereby extending BS ToA sensing capabilities to those devices. A BS could distinguish those, for example, by the ToAs themselves, the rough AoA of the incoming signals, or other factors [26]. RISs then are considered as anchors, with a decrease in the accuracy due to the additional reflection and the increased path length. We then use the same value for the time resolution and, by extension, the same $\sigma^2_{min,ToA}$ as for the BSs, but we do update the SNR of the measurement to reflect the accuracy degradation.

D. AoA sensing procedure

We model RIS devices with the capability to quickly reconfigure, able to follow and focus on the UE they are currently serving. This operating regime has the benefit of

[2]Neither of these assumptions is required formally nor computationally for our method: we exploit them to ease the exposition both in terms of formulas and graphical depictions.

offering precise angular information [27]. If we consider a sensor in the location $\mathbf{p}_s = (x_s, y_s, z_s)$ and a UE located at $\mathbf{p}_u = (x_u, y_u, z_u)$, an AoA measurement in horizontal coordinates translates into the following equations:

$$\theta_A = \text{atan2}(y_u - y_s, x_u - x_s), \tag{4}$$

$$\varphi_A = \arcsin(\frac{z}{\sqrt{x^2 + y^2 + z^2}}), \tag{5}$$

where the values φ_A and θ_A are the central values of angular measurements of the AoA in the local horizontal coordinate system of the sensor. These equations simply come from the definition of horizontal coordinates, where the elevation φ_A is measured from the horizontal plane XY or $Z = 0$, and the azimuth angle θ_A employs the atan2(y, x) function, which is defined specifically for this purpose as described in Section I. Since the statistical distribution of measurements is assumed unbiased, the central values of the measurement distributions are the true values.

We model the accuracy of the sensor in any given measurement as equal to its base accuracy, with a degradation factor dependent on the SNR [28]. If the sensor can achieve in optimal conditions a minimum value σ^2_{min}, we model the accuracy dependence with the SNR by dividing that value with a sigmoid function ranging between the value 0 at the minimum SNR (with a sensitivity threshold SNR_{min}), and 1 at the saturation SNR beyond which there are no noticeable improvements in signal quality, noted as SNR_{max}. Having chosen the logistic function, the sensor variance at the measurement σ^2_{AoA} yields

$$\sigma^2_{AoA} = \sigma^2_{min} \cdot \frac{1 + e^{k_{\text{SNR}}(\text{SNR}_m - \text{SNR}_c)}}{e^{k_{\text{SNR}}(\text{SNR}_m - \text{SNR}_c)}}, \tag{6}$$

$$\text{SNR}_c = \frac{\text{SNR}_{max} + \text{SNR}_{min}}{2}, \tag{7}$$

$$k_{\text{SNR}} = \frac{2 \cdot \ln(9)}{(\text{SNR}_{max} - \text{SNR}_{min})}, \tag{8}$$

where SNR_m represents the SNR obtained by the sensor at the measurement, and we choose the factors k_{SNR} and SNR_c to ensure the intended behavior of the model. What we intend with this model for σ^2_{AoA} is to encode the dependence of the sensor reliability with SNR. In Eq. (6) for high values of $\text{SNR}_m \geq \text{SNR}_{max}$ the sensor accuracy tends to the optimal value $\sigma^2_{AoA} \approx \sigma^2_{min}$, and for low values when $\text{SNR}_m < \text{SNR}_{min}$ the accuracy falls as $\sigma^2_{AoA} \gg \sigma^2_{min}$, with a smooth transition in between. This smooth transition is governed by the factors SNR_c and k_{SNR}: with the values chosen in Eq. (7) and Eq. (8) we have $\sigma^2_{AoA} = \sigma^2_{min}/0.9$ when $\text{SNR}_m = \text{SNR}_{max}$, representing near-optimal operation and $\sigma^2_{AoA} = \sigma^2_{min}/0.1$ when $\text{SNR}_m = \text{SNR}_{min}$, representing the absence of usable information in that circumstance.

We take the minimum variance σ^2_{min} of the measurements as the sensor resolution. The resolution in our model depends on the shape of the beam-pattern produced by the RIS, particularly the beam-width. Hence, the defining characteristic of the resolution of the RIS sensor will ultimately be the number and distribution of antenna elements on that device. To maintain a consistent form with the sensor Gaussian model detailed before, we model this beam-width as half the angular span between the two closest points to the beam centre that present a gain equal to $G_{BW} = G_{max} \cdot e^{-\frac{1}{2}}$, where G_{max} is the gain in the centre of the beam. As for BSs, we distinguish between fixed antennas, without a precoder or any other means to produce a configurable directional beam, and MIMO BSs, composed of an uniform rectangular arrays (URAs), uniform circular arrays (UCAs) or other similar means able to focus on the individual users at will, and hence producing angular information similar to that produced by RISs.

E. Shadowing and connection models

We consider a known environment, where we can evaluate whether a pair of points are linked through a LoS or not. We take into account each CS individually, and check the availability of LoS with all the points included in the area where UEs are expected to be distributed. The areas where this LoS condition is not fulfilled are noted as blockage areas and they are listed for each CS. Considering the propagation characteristics of the radio frequency bands where the RIS is an effective solution, non-line-of-sight (NLoS) links caused by reflections in the environment are excluded, as their power contribution is minimal under these conditions. This dichotomous behaviour is not only theoretically predicted but it has been also widely observed in simulations and field measurements [29], [30], providing further justification for our binary shadowing model. Having taken into account blockage with this mechanism, we can model remaining LoS connections with a simple, low-computational-complexity form of the Friis transmission equation:

$$P_r = P_t \cdot G_t \cdot G_r \cdot \frac{\lambda^2}{(4\pi d)^\gamma}, \tag{9}$$

where P_t and P_r are the transmitted and the received power, G_t and G_r are the transmitting and receiving gains, λ is the wavelength used for the transmission, γ is the path loss exponent, and d is the distance between the transmitting and the receiving antennas, all in metric units. While this simple model allows us to focus our work on the localization aspect, it is worth mention that the RL architecture we propose can be swiftly adapted to employ models with arbitrarily high detail, as long as they can provide a map of the chosen parameter to evaluate the communication quality (in our case the SNR), given a fixed infrastructure. For example, if a detailed and realistic simulator as the one described in [31] is available, our proposal can integrate its output to obtain much more refined solutions on the communications side, without any further adaptation or change needed on our proposal. These improvements on the simulation of individual scenarios could be expensive while not offering enough insight about the general case, and thus might be unfit for general research. Nevertheless the higher reliability of the data obtained from more detailed simulations can be the basis of a commercial exploitation of both the present work and [31].

IV. Statistical positioning quality

In this section we detail the fundamentals of our solution from a mathematical perspective. In particular, we consolidate our proposal premises and provide a solid guideline that allowed us to design in the next section our novel RIS-based ISAC framework.

A. Measurement projections

Given the heterogeneity of available measurements, we need to tailor all such information onto the same space. Hence, we map each one-dimensional measurement with its variance into the full-dimensional space of states for the UE, which includes 3 spatial dimensions (e.g., for the UE position, x_u, y_u, z_u) and one time dimension (e.g., for the UE internal clock deviation Δ_{t_u}) in the original form, or only 2 spatial dimensions (x_u, y_u) in the simplified form.

The result of the mapping is one analytical PDF for each measurement, assuming a Gaussian distribution of errors with the given standard deviation, individual to each measurement and dependent on the characteristics of the sensing device and measurement conditions. Since we assume the measurements to be normally distributed, all the information they carry is encoded in their central value μ and variance σ^2. Since the mapping between a pair of values (μ, σ^2) of any given measurement and their corresponding PDF is surjective we consider it a projection. With these PDFs we express the measurement *relative likelihood* of producing a value, given the full state definition. The PDF $P(\mathbf{s}|m)$ spanning over all the space of states $\mathbf{s}$ conditioned to a measurement real value m_r can be expressed as the following

$$P(\mathbf{s}|m_r) \propto P_s(\widehat{m}(\mathbf{s})|m_r), \tag{10}$$

where $P_s(\widehat{m}|m_r)$ is the relative likelihood of obtaining a measurement $\widehat{m}$ from a sensor given the real value m_r, and $\widehat{m}(\mathbf{s})$ is the measurement associated with state $\mathbf{s}$.

While both the mathematical derivations and the computational counterparts can be performed in the original $\mathbb{R}^4$ space, for simplicity of the exposition we show here the 2-dimensional PDFs where we assume the UEs are located in $z_u = z_c$ plane while discarding the UE internal clock deviation (i.e., $\Delta_{t_u} = 0$) and the AoA elevation measurements φ_A. As a measurement value can be produced by infinite points in the state space, we need more information to create the projection from $\mathbb{R}$ to $\mathbb{R}^n$. We assume that the state space is equiprobable [3], and hence all possible states producing the same measurement are equiprobable too.

B. Location probability distribution

Let us consider a sensor in the location $\mathbf{p}_s = (x_s, y_s, z_s)$ and a UE located at $\mathbf{p}_u = (x_u, y_u, z_u)$. If we derive the PDF encoding the spatial information from an AoA measurement $\widehat{\theta}$ using Eq. (4) and the Gaussian distribution, it yields

$$P_{AoA}(x,y) = C_{N,AoA} \cdot e^{\frac{-1}{2} \cdot \frac{D^2_{AoA}(x,y)}{\sigma^2_{AoA}}}, \tag{11}$$

$$D_{AoA}(x,y) = \widehat{\theta} - \theta_r, \tag{12}$$

$$\widehat{\theta} = \text{atan2}(y - y_s, x - x_s), \tag{13}$$

$$\theta_r = \text{atan2}(y_u - y_s, x_u - x_s), \tag{14}$$

$$C_{N,AoA} = \left(\int_{\mathbb{R}} \int_{\mathbb{R}} e^{\frac{-1}{2} \cdot \frac{D^2_{AoA}(x,y)}{\sigma^2_{ToA}}} \, dx \, dy \right)^{-1}, \tag{15}$$

[3]This assumption makes our analysis more tractable leading to an UE located in any position and have any internal clock deviation within the considered space with the same probability.

as depicted in Fig. 3 on the left-side hand in blue color. The intuition behind equations Eq. (11)-(15) is simple: for any given angular measurement $\widehat{\theta}$ of the user position, the user has a probability density $P_{AoA}(x, y)$ of being located in the point (x, y). To build this we simply consider that our angular measurements $\widehat{\theta}$ will have a normal distribution. Hence, as written before, we take the Gaussian distribution function, center it at θ_r following Eq. (4), and spread it accordingly to the expected sensor accuracy σ^2_{AoA} from Eq. (6). That leaves us with $\mathcal{N}(\theta_r, \sigma^2_{AoA})$, as seen in Eq. (12) with the only additional factor of the normalization constant $C_{N,AoA}$. Thanks to this factor, defined in Eq. (15), we normalize the total probability of $\int_{\mathbb{R}} \int_{\mathbb{R}} P_{AoA}(x, y) \; dx \; dy = 1$. Similarly to this, the PDF encoding the spatial information from an ToA measurement $\widehat{T_A}$ using Eq. (2) and the Gaussian distribution can be written as the following

$$P_{ToA}(x,y) = C_{N,ToA} \cdot e^{\frac{-1}{2} \cdot \frac{D^2_{ToA}(x,y)}{\sigma^2_{ToA}}}, \tag{16}$$

$$D_{ToA}(x,y) = (\widehat{T_A} - T_{A,r}), \tag{17}$$

$$\widehat{T_A} = \sqrt{(y - y_s)^2 + (x - x_s)^2}/c \tag{18}$$

$$T_{A,r} = \sqrt{(y_u - y_s)^2 + (x_u - x_s)^2}/c \tag{19}$$

$$C_{N,ToA} = \left(\int_{\mathbb{R}} \int_{\mathbb{R}} e^{\frac{-1}{2} \cdot \frac{D^2_{ToA}(x,y)}{\sigma^2_{ToA}}} \, dx \, dy \right)^{-1}, \tag{20}$$

depicted (unnormalised) in Fig. 3 in the center in green color. As per the previous group, Eqs. (16)-(20) allow us to compute the probability density $P_{ToA}(x, y)$ of the user being in a position (x, y), given this time a ToA measurement $\widehat{T_A}$. The construction of these equations to model the PDF for the ToA measurements is analogous to the previous case in Eqs. (11)-(15), in this case obtaining $\widehat{T_A}$ from Eq. (2) and $\sigma^2_{ToA}$from Eq. (3). We have again a normalised total probability of $\int_{\mathbb{R}} \int_{\mathbb{R}} P_{ToA}(x, y) \; dx \; dy = 1$ thanks to the factor $C_{N,ToA}$. As we consider the distributions of the internal errors of the sensors to be independent between devices, we can combine the available information by multiplying the PDFs as the following

$$P_R(x,y) = C_{N,R} \cdot \prod_i P_i(x,y), \tag{21}$$

and re-normalising the result using the variable $C_{N,R}$ to keep

$$\int_{\mathbb{R}} \int_{\mathbb{R}} \prod_i P_i(x,y) \; dx \; dy = 1, \tag{22}$$

which we depict for one ToA PDF and one AoA PDF in Fig. 3 on the right-hand side in red color. It is worth noting that, unlike the ToA PDF written in Eq. (16), the AoA PDF written in Eq. (11) is equal to zero. This is because as per Eq. (11) the normalization constant collapses to $C_{N,ToA} = 0$ since the double integral diverges, as it can be seen in the center of Fig. 3, where the (infinitely spanning) PDF is cut with a dashed line. Yet, when multiplied by other distributions as in Eq. (21), either angular as Eq. (11) itself (but with different values for $\mathbf{p}_s$ or $\mathbf{p}_u$), or radial as Eq. (16), and then renormalised as Eq. (22),

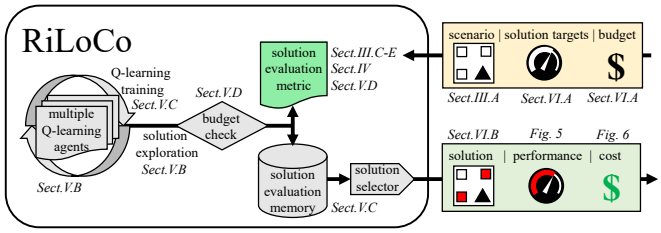

Fig. 4. Overview of RiLoCo building blocks.

its information is preserved and traduced to the final result. With this in mind, we can formulate the following theorem.

Theorem 1: For a given pair of sensor accuracies in AoA and ToA measurements $(\sigma_{ToA}, \sigma_{AoA})$, and with the objective of maximizing localization accuracy, RIS behave as an optimally placed pair of sensors.

Sketch of Proof: Consider a PDF f, which encodes a set of measurements ε attaining a localization precision of P. Let λ_i be the eigenvalues of $\mathcal{I}(\mathbf{q})$, the Fisher information matrix of f, which provide an upper bound for P, i.e. $\max_{\varepsilon}\{P\} \leq \lambda_i$. Following Eq. (11) and Eq. (16), we define two different PDFs, $f_{ToA} : \mathbb{R}^4 \rightarrow \mathbb{R}$ and $f_{AoA} : \mathbb{R}^3 \rightarrow \mathbb{R}$, describing ToA and AoA measurements respectively. We denote by $f_R(x, y, z, t)$ the properly normalized PDF obtained by the product of f_{ToA} and f_{AoA}. Lastly, assume that both original measurements are obtained from two sensors whose positions, given by $\vec{p}_{s,ToA}, \vec{p}_{s,AoA}$, are pair-wise independent. Then, $\max_{\vec{p}_{s,ToA},\vec{p}_{s,AoA}} \lambda_i = \lambda_{\vec{p}_{s,ToA}=\vec{p}_{s,AoA}}$, since the eigenvectors associated to the largest eigenvalues of the Fisher information matrices of f_{ToA} and f_{AoA} need to be spatially orthogonal. □

We provide an extensive description of our framework by shedding the light on the main novel building blocks and their corresponding interactions within a classical propagation environment.

C. Problem definition

We consider an area $\mathcal{A}_u$ with an extension A_u where we aim at providing communication and localization services to a homogeneous distribution of UEs. We have N predefined CS available for the installation of RISs or BSs. We define the deployment vector as $\mathbf{d} = \{d_i\}^N$, with d_i being the device installed the i-th CS. We have a maximum budget constraint β such that the overall installation cost is within such a budget and a function $\kappa(\mathbf{d})$ that economically evaluates any deployment solution. This budget can represent economical means or energy consumption, in order to optimize the energy efficiency of the infrastructure. We have a set of devices $\mathcal{V} = \{0, 1, ..., V\}^N$, each with a set of characteristics, available to be installed in each CS, and including an entry to represent the absence of any installed device. We use two initial metrics in our proposal. The localization metric is denoted as $L(\mathbf{d}, x_u, y_u)$, and the communication performance metric is denoted as $C(\mathbf{d}, x_u, y_u)$. We detail both metrics later in this section, subsection *D. Performance metrics*, with their individual descriptions highlighted in bold as **Communication** and **Localization**. We define the joint metric as

$$M(\mathbf{d}, x_u, y_u) = \alpha\ L(\mathbf{d}, x_u, y_u) + (1-\alpha)\ C(\mathbf{d}, x_u, y_u), \quad (23)$$

where α marks the balance between localization and communications. With $\alpha = 1$ the joint metric attends only to localization and $\alpha = 0$ attends only to communications. Now, we can formulate our optimization problem as the following

Problem 1 (Joint Sensing and Communications Optimization):

$$\max_{\mathbf{d}} \quad \frac{1}{A_u} \iint_{\mathcal{A}_u} M(\mathbf{d}, x_u, y_u)\, dx\ dy \quad (24a)$$

$$\text{subject to} \quad \kappa(\mathbf{d}) \leq \beta, \quad (24b)$$

$$\mathbf{d} \in \{0, 1, ..., V\}^N. \quad (24c)$$

Note that, in our formulation the budget observed in the constrain Eq. (24b) can be also easily incorporated in our metric Eq. (23) as a divisor, allowing to not only produce solutions optimizing the performance per watt, or abiding a capital expenditure ceiling, but also to search for the best return of investment (ROI), among other energy-efficiency and economically-guided solutions, as later shown and detailed in Section VI.

D. RiLoCo Architecture

Overview. The existing formulations of the deployment problem are, even without taking RISs into consideration, NP-Hard [32]. Hence, finding solutions that cover novel situations and paradigms in a practical, feasible way is an open problem. Nevertheless, in this paper the solving algorithm itself occupies a secondary position, as the core of the presented work is the metric that we build in this section, in the subsection *D. Performance metrics*. We advance the notation of our final metric in Eq. (23). We use this notation to condense the deployment problem in the form of Problem 1. In subsection *C. RiLoCo in Action* we explain how we intend to use our metric to guide a straightforward solution search method to obtain good quality solutions, and in subsection *D. Performance metrics* we explain how our metric conforms a direct representation of the communications and sensing capabilities of an infrastructure. With this objective, and to make Problem 1 tractable, we leverage on multiple Q-learning RL agents to explore the feasible solutions, as shown in Fig. 4 and Alg. 2. The foundations of Q-learning we employ for this are explained in this subsection, in the **Reinforcement learning** part. To employ this technique, we translate the set of possible solutions to Problem 1 into a space of states. As explained in the **Reinforcement learning** part, the Q-learning agent will then navigate this space of states by taking actions. Additionally, to guide the agent training we devise our localization and communication metrics, and a way to combine them in different degrees to produce a single, final metric that serves as feedback, evaluating the quality of each solution found, and by extension the quality of each action taken by the agent. These metrics are built by leveraging two key concepts: i) the Fisher information and ii) the Logistic function.

We would like to emphasize that Q-learning is chosen for our architecture because the discrete nature of the problem

formulation makes it adequate for our situation, for its adaptability and guaranteed convergence, but also because its simplicity. More complex solutions employing deep reinforcement learning could be employed, having shown in previous works as [8] ample capacity to solve the deployment problem and interesting capabilities (as better scalability). However they come at a cost that we expressly avoid: a more convoluted architecture.

Reinforcement learning. RIS deployment problems often are, when no simplifications or approximations are employed, NP-Hard [33]. Traditional tools to create deployment methods rely on relaxations and adaptations of the problem to make it tractable, or solutions are found using approximate methods. While these approaches work, the alternative option of using machine learning (ML) methods, and in particular RL, allows us to tackle the problem with as much realism as the simulator employed allows for [8], [33]. To showcase how our metric can lead a simple RL agent to solve the problem at hand, we opt for Q-learning. While there is no dynamic situation for the agent to handle, we create an environment where the agent continuously alters the deployment solution to maximize its performance.

Q-learning is a model-free RL algorithm. It learns to optimize the reward obtained by actions in an environment by elaborating a table (called the Q-table) with one entry for every possible state and action, based on the past rewards received. For RiLoCo we build on the traditional Q-learning architecture, as introduced originally in [34], [35], and schematically presented in Algoritm 1. On it, we denote the discount factor by γ, which governs the ability of the agent to take future expected rewards into consideration, instead of following a greedy approach.

Algorithm 1 Q-Learning Algorithm

Initialize Q-table arbitrarily (e.g., with zeros) for all state-action pairs
Set parameters: learning rate α, discount factor γ, exploration rate ϵ
for each episode **do**
  Initialize state s_0
  for each step in the episode **do**
    Choose action a_t based on the exploration-exploitation strategy
    Take action a_t, observe reward r_{t+1}, and transition to next state s_{t+1}
    Update Q-table:

$$Q(s_t, a_t) \leftarrow (1-\alpha)\cdot Q(s_t, a_t) + \alpha\left(r_{t+1} + \gamma \max_a Q(s_{t+1}, a)\right)$$

    Set $s_t \leftarrow s_{t+1}$
  end for
end for

V. RiLoCo framework

The rationale behind our choice relies on the fact that Q-learning is proven to converge to the optimal solution [35].

Algorithm 2 Overview of RiLoCo in pseudocode form.

Initialize solution evaluation memory
while Total training time not exhausted **do**
  Instantiate Q-learning agent
  while Individual training time not exhausted **do**
    Train Q-learning agent (Alg.1)
    if Solution state s_t or s_{t+1} exist in memory **then**
      Fetch solution evaluation from memory
    else
      Evaluate $\frac{1}{A_u}\iint_{\mathcal{A}_u} M(\mathbf{d}, x_u, y_u)\, dx\ dy$
      Store solution evaluation in memory
    end if
    if Budget $\kappa(\mathbf{d}) > \beta$ **then**
      $s_{t+1} \leftarrow s_t$
      $r_{t+1} \leftarrow 0$
    end if
    Update Q-table (Alg.1)
  end while
end while
Select solution with best performance from solution evaluation memory within budget $\kappa(\mathbf{d}) \leq \beta$

In our specific problem we encode each possible deployment solution as an state s. Any deployment solution can be encoded in a vector with n positions, one for each CS in our scenario. For each position of the vector (each CS) we have m mutually exclusive deployment choices: leaving the site empty and $m-1$ different options of devices to install. That way, the total number of possible solutions is m^n. Since the set of possible solutions is finite, we establish a bijective relation between each solution vector and a state number. The set of actions we define for the agent to navigate in this space is composed of $2 \cdot n$ options: for each CS in the vector the agent can choose to increase or decrease the index of the device installed on that site, being able to empty it. This leaves us with a Q-table with m^n rows and $2 \cdot n$ columns.

For the training, we reward each action of the agent with the difference in the metric valuation between the resulting state and the original state of each action. That way, the total-reward convergence of the Q-learning agent [35] warrants the convergence to the optimal solution. We discard the actions that would produce a state which would exceed the given budget. In those cases, we do not change the current state and we feed the agent a zero reward.

The balance between exploitation and exploration in the training action decisions follows a linear evolution, from all-exploration at the beginning to all-exploitation at the end.

Logistic function. We face several challenges while building both our proposed deployment method RiLoCo, its system model and formal framework. One particularly recurring and central obstacle is the adequate handling of extreme cases in both the localization and communication metrics. Such extreme cases may range between areas (e.g., very close to a BSs and RISs) that attain location and communication performances several orders of magnitude above others, and far-away areas where services simply do not reach a usable threshold. To encode increases and decreases in already

outlying values while ensuring tractability of the resulting performance indicators using data driven and ML tools, we use sigmoid functions as a way to preprocess raw values. Among sigmoid functions, we opt for the logistic function to exploit the convenient properties it shows (e.g., symmetry) to make mathematical tractability easier. We then build the metrics here presented using said logistic function, advanced in Section III as part of Eq. (6), with the following form

$$g(x) = \frac{e^{k(x-x_0)}}{1+e^{k(x-x_0)}}. \tag{25}$$

By carefully selecting the parameters k, x_0 in the function, we can choose the thresholds between negligible function values and saturated output. This behaviour leads to an optimal use of both communication and localization services as described hereafter. Other functions could serve in these objectives, e.g., saturating the values with maximum and minimum thresholds τ_{max} and τ_{min}, as $g_s(x) = \max(\tau_{min}, \min(\tau_{max}, x))$, but the logistic function favours its usage for ML training efforts being smooth for all $x \in \mathbb{R}$.

Fisher information. The Fisher information is a metric to evaluate the amount of information we can obtain about an unknown parameter q from an observable random variable Z. In our case, we want to know how much information we can obtain about the $\mathbb{R}^4$ parameter comprising the original UE position and internal clock deviation $\mathbf{q} = (\mathbf{p}_u, \Delta_{t_u}) = (x_u, y_u, z_u, \Delta_{t_u})$ from all our sensor measurements combined, translated to a single PDF.

When considering a parameter $\mathbf{q}$, estimated based on an observation vector Z, which has a PDF $P(Z)$, the Fisher information takes the form of a symmetric square matrix with the same dimension as $\mathbf{q}$, with the element in the i-th row and j-th column defined [36] as

$$\mathcal{I}(\mathbf{q})_{i,j} = \mathbb{E}\left[\left(\frac{\partial \ln P(Z)}{\partial q_i}\right)\left(\frac{\partial \ln P(Z)}{\partial q_j}\right)\right]. \tag{26}$$

Our specific application of the Fisher information is detailed later in this section, subsection *D. Performance metrics*, in the **Localization** part. Essentially, we employ the statistical distribution of measurements we derive from the sensor models we propose to compute the Fisher information. Numerically, this direct application implies the computation of the PDF in a fine grid, obtaining the logarithm in each point of the grid, and the exhaustive computation of Eq. (26).

The CRB states that the Fisher information is the upper bound of the precision of any unbiased estimator. Therefore, we use the Fisher information to evaluate each sensor network. Founding RiLoCo in the Fisher information we can avoid the assumption of a particular localization framework, and hence avoid the evaluation of the localization performance using that method. Instead, we evaluate the amount of information available from our measurement set about the UE coordinates.

A. RiLoCo in Action

We propose an iterative method for installing and configuring RISs to jointly improve localization and communication performances, namely RiLoCo. It works in cycles, where we take random deployments and alter them step by step using a RL agent. The iterative deployment process using the agent does not follow a greedy policy: by rewarding the agent with the difference between the performance metrics of the current and the previous scenario, we train the agent to maximize the performance of the final solution, not the intermediate steps. While training the agents is a fast process, the different deployment evaluations take a lot of computational effort. We obtain the best results by storing these results in a solution evaluation memory (see Fig. 4 and Alg. 2), and run over it a series of Q-learning agents, which can request any solution evaluation, buffered or not. By running agents that focus their starting random deployments on the best scenarios previously found, we are able to introduce a refining phase for the solutions, which in practice greatly reduces the time needed to reach a certain level of solution performance.

Solution steps. The high-level descriptions of the solution stages are i) considering a random scenario, we evaluate it with the localization and communication metric described in Section V-B, Eq. (32), ii) altering the scenario using a RL agent which attempts to maximize said metric, and iii) deciding if the process continues (going back to the step i) based on the quality of the current solution and the improvement achieved in the last iterations [4]. The training is simply exploited as a searching method, refining the solution by attempting a balanced mix of informed changes based on the knowledge acquired from the quality of past probed solutions and trying unexplored options. Hence, the success of this algorithm at finding a good deployment is independent of the success of the training phase of the agent. This gives us the capability of tackling the problem in situations where a traditional approach would fail, either for the requirement of much larger computational resources or because of the simple unfeasibility of it.

Proposed localization framework. We make heavy use of the Fisher information to make our results independent from the particular localization method used to infer the UE location from the sensor measurements. Nevertheless, the same ideas of representing one-dimensional measurements in the full-dimensional state space where we want to locate the user can be exploited for a localization that takes into account all information at hand. In equations Eq. (11), (16) and related ones the PDF represents the probability of *obtaining* a measurement *given* a true state of the user and the sensor accuracy, but we can use it to create a localization method: given a measurement and the sensor accuracy, we can reconstruct the PDF of the true UE locations, and combine all sensors information in the same manner described in Eq. (21).

With this method we are able to integrate NLoS readings if the PDFs are built taking it into account, for example by changing the simple Gaussian distribution of the measurements to the convolution of that Gaussian and a flat probability of producing a NLoS reading (a *bias*) in a range, as proposed in Eq. (9)/Fig. 2 of [37]. This method, while being computationally heavy, can integrate measurements from

[4]Note that the solving algorithm is built as a ML training process for a RL agent, but the successful production of a deploying-capable RL agent is neither necessary nor the objective of the solution process.

wildly heterogeneous sources of information, and seamlessly combines all the measurements from the sensors without loss of information neither on the measured value nor in the variances and covariances of the final distribution. While we do not showcase the results of this method in the present paper, we have used it to introduce an additional measurement improvement, explained in the following.

Improved location probability distributions. Building on the former localization framework, we propose a method to discard NLoS readings for localization. By including an additional step in the localization accuracy measurement we can filter the NLoS measurements. When we have obtained the combined PDF of all the available measurements, the overlap of that total and the individual measurements, defined as the scalar product between the two functions $\langle u, v\rangle = \iint_{\mathbb{R}^2} u(x,y)v(x,y)\,dx\;dy$ can serve as a guide to identify the (probable) NLoS measurements. By discarding those measurements, either AoA or ToA or others that do not overlap with the estimation, we can refine the initial measurement. The convolution of the original measurements with the random NLoS bias can be discarded, and the original process can be reproduced with the LoS measurements, to produce a more precise result.

B. Performance metrics

The core process within RiLoCo is to explore the solution space while attempting to maximize our metric (Eq. (23)). To guide the exploration we reward each action of the agent with the difference between the valuation of the initial scenario (before the action was taken) and the resulting scenario (after the action is taken). The quality of the solutions found will always have the quality of the metric they rely on as a limiting factor. This makes the performance metrics a point of major importance in our proposal.

To jointly optimize localization and communication performances, we build a joint metric by combining two individual metrics, as shown in Eq. (23), both adequately evaluating the solutions.

Communication. Our metric $C(\mathbf{d}, x_u, y_u)$ for the communication performance can be considered the simplest of the two and it is based on the SNR that each deployment $\mathbf{d}$ can provide for a fictitious UE placed on coordinates x_u, y_u on the considered surface. To obtain the SNR in that point we first model the present RISs as signal sources. We compute the different amounts of power received by one antenna element of the i-th RIS (which contains ne_i elements) and choose the BS that shows the highest power transmission to the RIS: $P_{ei} = \underset{j}{max}(P_{j-i})$, where P_{ei} is the maximum power received by an antenna element of the i-th RIS and P_{j-i} is the power received by one antenna element of the i-th RIS from the j-th BS. We multiply that power by the number of elements in the RIS to obtain the maximum signal power the RIS is able to reflect, $P_i = P_{ei} \cdot ne_i$. Knowing the distribution of elements of the RIS and treating it as a URA [38] able to focus on each direction, we compute the maximum attainable gain by the RIS for every AoA of the signals, $G_i(\theta)$. For the UE in location (x_u, y_u) the angle between the RIS and the user is $\theta_{i-u} = \text{atan2}(y_u - y_{si}, x_u - x_{si}) - \theta_i$, where (x_{si}, y_{si}) is the location of the i-th RIS and θ_i is its orientation with respect to the reference direction. Having the maximum power that the RIS can reflect P_i and the gain it shows in the direction of the UE, and assuming an isotropic antenna in the UE, we simply apply Eq. (9) to obtain the received power by the UE [38]. We obtain the SNR by dividing this amount by a background noise level. Considering the values of the SNR received from all the present signal emitters (both RISs and BSs), we use the highest one as the SNR available for the UE $\text{SNR}_u = \underset{k}{max}(\text{SNR}_{u,k})$, where $\text{SNR}_{u,k}$ is the SNR received by a UE at the position x_u, y_u from the k-th signal emitter. To turn this value into a score between 0 and 1, we apply the logistic function Eq. (25). To choose the parameters k, x_0 we take as milestone values the expected minimum threshold at which mobile radio connectivity can be provided in beyond fifth-generation (B5G), and the expected maximum threshold beyond which no noticeable further improvement is obtained, i.e., SNR_{min} and SNR_{max}, accordingly. We finally have the following

$$C(\mathbf{d}, x_u, y_u) = \frac{e^{k_{\text{SNR}}(\text{SNR}_u - \text{SNR}_c)}}{1 + e^{k_{\text{SNR}}(\text{SNR}_u - \text{SNR}_c)}}. \tag{27}$$

Since we want this function to return a value of 0.9 when $\text{SNR}_u = \text{SNR}_{max}$, and by 0.1 when $\text{SNR}_u = \text{SNR}_{min}$, we pick k_{SNR} and SNR_c as per Eq. (8) and Eq. (7).

To turn this metric $C(\mathbf{d}, x_u, y_u)$, which evaluates a single point, into the scenario-wide metric $C_S(\mathbf{d})$ that assesses the deployment $\mathbf{d}$ over the considered area $\mathcal{A}_u$, we simply take and normalize the integral over $\mathcal{A}_u$ by averaging the value of the metric as follows

$$C_S(\mathbf{d}) = \frac{1}{A_u} \iint_{\mathcal{A}_u} C(\mathbf{d}, x_u, y_u)\,dx\;dy. \tag{28}$$

Localization. Our metric $L(\mathbf{d}, x_u, y_u)$ for the localization performance is based on the Fisher information of the PDF of the UE location measurements. We first consider all the possible measurement distributions for a fictitious UE placed on coordinates (x_u, y_u) on the considered surface, from all the available sensors in the scenario. We then obtain the measured UE position PDFs. These take the form of Eq. (11) for AoA measurements and Eq. (16) for ToA measurements. We condense them in a single PDF $P_R(x, y)$ following Eq. (21): this digests all the positional available information from the sensors in the scenario. Hence, we compute its Fisher information matrix, with the parameter $\mathbf{q} = (x, y)$ in Eq. (26) where $q_1 = x$ and $q_2 = y$, it yields

$$\mathcal{I}(\mathbf{q}|x_u, y_u)_{i,j} = \mathbb{E}\left[\left(\frac{\partial \ln P_R(x,y)}{\partial q_i}\right)\left(\frac{\partial \ln P_R(x,y)}{\partial q_j}\right)\right]. \tag{29}$$

In practice, as introduced in the **Fisher information** part of subsection *B. RiLoCo Architecture*, we solve this equation numerically. To do this, we compute $P_R(x, y)$ in a fine grid, obtain the logarithm of each value of the grid, make the numerical derivatives in $q_1 = x$ and $q_2 = y$, and obtain the expected value in the formula through the product with the original distribution $P_R(x, y)$ and numerical integration.

The Fisher information matrix of $P_R(x, y)$ provides 3 distinct values, $\mathcal{I}(\mathbf{q})_{1,1}$, $\mathcal{I}(\mathbf{q})_{2,2}$ and $\mathcal{I}(\mathbf{q})_{1,2} = \mathcal{I}(\mathbf{q})_{2,1}$. Since the localization performance is ultimately limited by the minimum localization accuracy in any direction, and we want to make our metric independent under rotations, we summarize the information matrix into a single value, taking the smallest of the eigenvalues of the matrix. We diagonalise the Fisher information matrix as $\mathcal{I}(x, y) = P\mathcal{I}_D P^{-1}$. We take the smallest value of the diagonal in the new base as the minimum spatial information in any direction, $\mathcal{I}_{min} = min\{ev_1(\mathcal{I}), ev_2(\mathcal{I})\}$.

Leveraging on the CRB, we compute the inverse of the square root to obtain the minimum bound of the standard deviation any unbiased position estimator would be able to achieve when locating the UEs in the direction of greatest uncertainty, $\sigma_{\text{CRB},u} = 1/\sqrt{\mathcal{I}_{min}}$, using the deployment of sensing infrastructure $\mathbf{d}$ initially given. Localization services are usually required to meet a sharp threshold. Increased accuracy above this threshold brings no additional benefit, and performance falling short of it renders the service valueless. To model this, given a localization accuracy threshold σ_{TH}, we resort again to the logistic function Eq. (25) (but inverted, since we want to reward low $\sigma_{\text{CRB},u}$ values) to obtain our final metric as the following:

$$L(\mathbf{d}, x_u, y_u) = \frac{1 + e^{k_{\text{CRB}}(\sigma_{\text{CRB},u} - \sigma_{\text{TH}})}}{e^{k_{\text{CRB}}(\sigma_{\text{CRB},u} - \sigma_{\text{TH}})}}. \tag{30}$$

In this case, the threshold value σ_{TH} has to be chosen as the localization accuracy that our service aims to provide, and k_{CRB} needs to be high enough to render the logistic function a soft step, but keeping a slope that enables ML methods. When using meters for the $\sigma_{\text{CRB},u}$ values, a value of $k_{\text{CRB}} = 10 \cdot ln9 \approx 21.97$ provides a change from 0.1 to 0.9 when going from $\sigma_{\text{TH}} - 10cm$ to $\sigma_{\text{TH}} + 10cm$ and a central slope of $5m^{-1}$. In our numerical evaluation we use a $\sigma_{\text{TH}} = 1$m.

As with the previous metric, to turn this metric $L(\mathbf{d}, x_u, y_u)$, which evaluates a single point, into the scenario-wide metric $L_S(\mathbf{d})$, which evaluates the deployment $\mathbf{d}$ over the whole area $\mathcal{A}_u$, we simply take and normalize the integral over $\mathcal{A}_u$, averaging the value of the metric:

$$L_S(\mathbf{d}) = \frac{1}{A_u} \iint_{\mathcal{A}_u} L(\mathbf{d}, x_u, y_u)\, dx\ dy. \tag{31}$$

Naturally, we extend Eq. (23) to the scenario-wide metrics Eq. (28) and Eq. (31) as follows:

$$M_S(\mathbf{d}) = \alpha\ L_S(\mathbf{d}) + (1 - \alpha)\ C_S(\mathbf{d}). \tag{32}$$

Budget. An additional external factor we take into account for deployments is the budget constraint. This energy or cost limitation takes the form of a set of rules that can evaluate the resources required by any proposed solution and rule it viable or not viable. The variables included in this budget evaluation are the total available budget and any relevant RIS or BS characteristics, such as size, number of elements, sensing capabilities, or location. The budget rule acts simply by discarding unfeasible solutions, so it does not need to follow any linearity or derivability conditions, simply to deem a deployment doable or not.

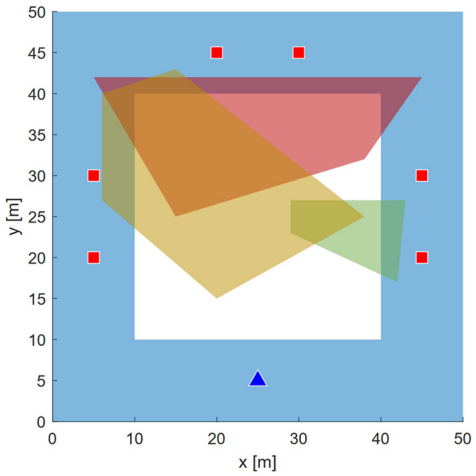

Fig. 5. Example of a base scenario, with CSs, one BS and shadowing areas.

VI. Numerical Evaluation

We evaluate RiLoCo by means of exhaustive numerical simulations using a commercial tool, namely MATLAB.

A. Methodology and simulation specifics

We use discrete probability distributions as homogeneous samplings of the corresponding continuous form: the value of the individual samples does not change with the spatial frequency of the sampling. If a two-dimensional PDF P is sampled within an area in the points $\mathbf{p}_{i,j} = (x_i, y_j)$ where $\Delta_x = x_i - x_{i-1}\ \forall i > 1$ and $\Delta_y = y_j - y_{j-1}\ \forall j > 1$, it is normalised as $\sum_{i,j} P_{i,j} \cdot \Delta_x \cdot \Delta_y = 1$, where $P_{i,j}$ is the i, j-th sample of the PDF P and Δ_x, Δ_y are the spatial granularity of the sampling. We use a spatial granularity of 1m.

In addition, we assume the values for SNR_{min} and SNR_{max} according to current technological typical values at 0dB and 20dB respectively, based on [39]. As per localization, we take an accuracy threshold of $\sigma_{\text{TH}} = 1$m as a conservative reference for near future networks [40], and $\gamma = 2$.

The scenario for the numerical evaluation is depicted in Fig. 5. On it, we can see, from a zenithal view, the complete playground forming a $50m \times 50m$ square in the $z = 0m$ plane, from the origin to the $x = 50m, y = 50m$ point. In the central area, centered in the scenario, measuring $30m \times 30m$, and painted white in Fig. 5, we distribute homogeneously the UEs. We use these UEs to probe the capabilities of the possible deployments of devices. This UE area goes from $x = 10m, y = 10m$ to $x = 40m, y = 40m$. In the bottom of Fig. 5 we can see where we place the BS we consider in the simulation, at $x = 25m, y = 5m$. In this example we consider 6 possible places where an RIS can be placed, i.e., the CSs, and we represent them as red squares. The shaded regions in red, mustard, and green represent the shadow areas that affect the BS and CSs. The moderate number of CSs allows us to realistically simulate a medium-sized scenario while avoiding the steep growth of computational costs associated with the NP-Hardness of the problem [32]. We operate at a frequency of $f = 24$GHz with a bandwidth of $B = 100$MHz for the ToA measurements. We assume a background noise of -80dBm whereas we have two available deployable RIS models, one

with 40×10 elements and one with 80×20. The budget rule values these at $\beta_1 = 100\$$ and $\beta_2 = 400\$$ respectively with an install cost of $\beta_{ins} = 200\$$. We consider a total budget $\beta_{tot} = 1200\$$.[5]

B. Results

The Q-learning agents, using the buffered solution evaluation and the differential reward, are able to find outstanding results[6]. Traditional solutions to the challenge of deploying communication infrastructure typically focus solely on communication metrics. In our study, we obtain similar solutions to serve as benchmarks by running our solver with communication metrics as the only form of feedback, i.e. Eq. (32) with $\alpha = 0$. On the other hand, the—more scarce—sensor deployments only take into consideration the accuracy of their combined information. We can observe how the best results obtained by looking only at one specific metric according to Eq. (32) (with $\alpha = 0$ in case of communication-oriented metric and $\alpha = 1$ for localization-oriented metric), often show very poor performance towards the neglected metric, as suggested by Fig. 6. In other words, the trade-off of these two metrics is made apparent in said Fig. 6. It is clear that a single-objective solution search can result in a slight improvement in their metric of focus while significantly degrading the quality of the deployment in the other.

Indeed, the necessity of a proper localization metric can be further noticed in Fig. 7: Solutions with reasonable communication performance often lead to unacceptable localization accuracy. Furthermore, the usage of the combined metric has proven highly successful: the corresponding solutions usually exhibit at the same time almost-optimal performance in both communication and localization (Fig. 6). Such solutions often remain hidden when we consider communication or localization-oriented performance alone.

In Fig. 7 we can observe one of the side results of the solution buffer: we can analyze the best solutions when different budget rules are applied. This breakdown of the results can greatly help the determination of optimized deployments in infrastructure within economic or energy efficiency considerations. In the example, we can see how a budget limit β of half of the initially allocated amount can readily achieve almost the same performance as the best solution with the full resource expenditure. We can also compare deployments against the current infrastructure performance in both metrics used as baselines. Our presented solution can be easily expanded if provided with a realistic translation able to map performance gains to economic gains, taking into account projected network usage given additional coverage or throughput, projected market share under the improved performance, etc. With such information, the metrics and the solution that we propose in this paper can easily guide future deployments economic and energy efficiency terms, focused either on the optimization of the benefits obtained per consumed watt or the maximum ROI.

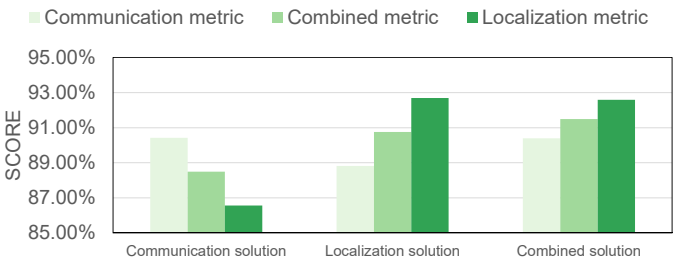

Fig. 6. Performance comparison of different solutions: communications-focused (legacy), localization-focused, and combined ISAC (RiLoCo) guided.

As explained in Sec. V, the core of our proposal, RiLoCo, is on the metrics used to guide the solution-finding process. We use Fig. 8 to show at the same time the aspect of these metrics, and the differences between the different deployments we find using RiLoCo. In Fig. 8 we have three columns (one for each solution found for the scenario) and two rows (one for each metric). The three solutions are those already represented in Fig. 6 and Fig. 7: one that maximizes the communication metric, one that maximizes the localization metric, and the last one, maximizing the combined metric. Regarding the rows, we devote the upper one to the localization metric: in particular, we plot $\sigma_{\text{CRB},u} = 1/\sqrt{\mathcal{I}_{min}}$, the core of the localization metric. This represents the maximum achievable accuracy that each deployment solution is able to reach across the simulated scenario $\mathcal{A}_u$. This value is chosen for its intuitive meaning: To use it, as detailed in Sec. V, we map it to the $(0, 1)$ interval with the logistic function to form Eq. (30), and then integrate it across the area $\mathcal{A}_u$ to obtain Eq. (31). In the lower row we represent the coverage metric. Similarly, we represent SNR, which process using the logistic function to obtain Eq. (27) and integrate in Eq. (28). With this we illustrate at the same time two important parts of this work: on one hand, how the localization metric measures the accuracy achievable by combining the information available from present sensors, and on the other hand, how we combine that with the communications metric. Fig. 8 serves to visualize how the best solutions found to satisfy each metric can diverge in shape.

In Fig. 8 we can see too where within $\mathcal{A}_u$ we meet or fall short of our communication and localization targets. We can see how the deployment greatly enhances communication in all cases, almost eliminating dead zones, exhibiting a minimum SNR of ≈ 7.5dB, way above $\text{SNR}_{min} = 0$dB. Notably, the experienced localization accuracy, not available at a single-BS infrastructure, hits the demanding performance of 1m in large spatial extensions in the localization and combined solutions.

While a formal, thorough complexity analysis would be impractical due to the random nature of Q-learning, we have observed that the time t_c needed for the convergence of the

[5]For the RIS cost estimation we take into account the number of elements of the considered devices, the frequency of $f = 24$GHz, an inter-element distance of $\lambda/2$, current costs for large batches of large-scale printed circuit boards, and an estimation of the costs per element in existing prototypes. Nevertheless, the exact values are not central to our results, as the relevant feature we want to showcase is the capability of RiLoCo to work within a budget and to produce different solutions within that constraint. As specified before, the budget constrain can represent a limit in either energy consumption, economic cost, or both.

[6]We have empirically tested that RiLoCo results are near-optimal by comparing them with the results from exhaustive searches; however, due to the complexity of the problem, this comparison could only be executed for very limited instances of the original problem.

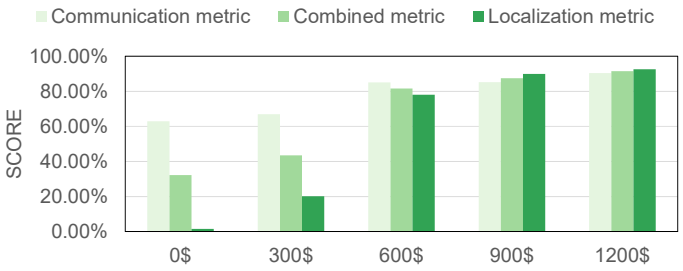

Fig. 7. Performance comparison of solutions for different budgets.

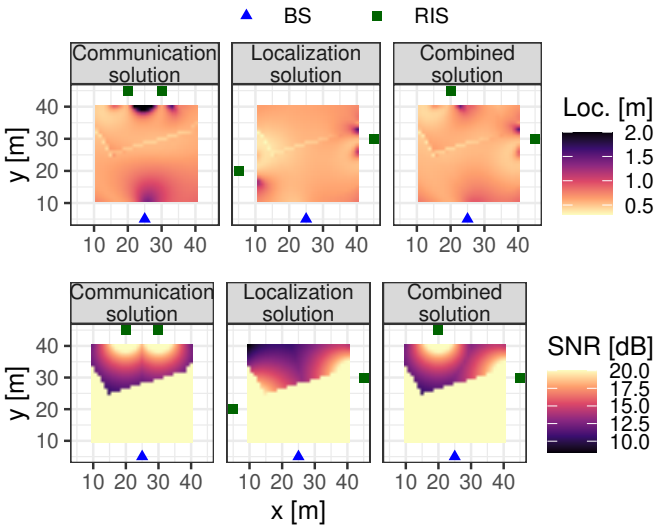

Fig. 8. Localization and communication metric maps of the communications-focused, localization-focused, and combined metric-guided solutions.

solutions shows an approximate linear dependence with the maximum number of devices that can be installed for the allocated budget $t_c \sim 1800s \cdot \beta_{tot}/(\min_i(\beta_i) + \beta_{ins})$, for an approximate time of two hours in the runs detailed here, running on a personal computer equipped with a 12th Gen Intel® Core™ i7 and 32 GBytes of random access memory.

VII. Conclusions

This paper highlights the transformative potential of RIS in powering the sensing and communications revolution within the context of 6G networks. RISs offer unprecedented opportunities to enhance the performance and efficiency of future wireless networks, enabling a new use of connectivity and data exchange. In particular, we have studied how both localization and communication-oriented design can coexist in wireless networks by proposing an evaluation metric. Based on it, we pioneer RiLoCo, a novel framework that optimally designs RIS placement, achieving with them an outstanding, next-gen localization performance attaining sub-meter accuracy, improving over traditional approaches focused either on communications-only or accuracy-only metrics.

Obtained results and the demonstrated flexibility and adaptability of our novel architecture will open the door for future work, for example, revolving around the usage of ray tracing simulation tools or more in-depth economic conditions for real-life commercial applications.

VIII. Acknowledgement

This work was partially supported by the SNS JU Horizon Europe project under Grant Agreement 101139130 (6G-DISAC), by the SNS JU Horizon Europe Project under Grant Agreement No. 101192521 (MultiX), by the SNS JU Horizon Europe Project under Grant Agreement No. 101139161 (INSTINCT), by the European Union under the Italian National Recovery and Resilience Plan (NRRP) of NextGenerationEU, partnership on "Telecommunications of the Future" (PE00000001 - program "RESTART") – CUP E63C22002040007, by the Spanish Ministry of Economic Affairs and Digital Transformation and the European Union – NextGeneration EU, in the framework of the Recovery Plan, Transformation and Resilience (PRTR) (Call UNICO I+D 5G 2021, ref. number TSI-063000-2021-6), and is part of the Grant 6G-DIFERENTE (CER-20231018) funded by CDTI and by the European Union NextGenerationEU/PRTR within the call "Ayudas Cervera para Centros Tecnológicos 2023"

References

[1] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 6, pp. 1728–1767, 2022.

[2] F. Liu, C. Masouros, A. P. Petropulu, H. Griffiths, and L. Hanzo, "Joint radar and communication design: Applications, state-of-the-art, and the road ahead," *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3834–3862, 2020.

[3] J. A. Zhang, M. L. Rahman, K. Wu, X. Huang, Y. J. Guo, S. Chen, and J. Yuan, "Enabling joint communication and radar sensing in mobile networks—a survey," *IEEE Communications Surveys & Tutorials*, vol. 24, no. 1, pp. 306–345, 2022.

[4] F. K. Gruber and E. A. Marengo, "New aspects of electromagnetic information theory for wireless and antenna systems," *IEEE Transactions on Antennas and Propagation*, vol. 56, no. 11, pp. 3470–3484, 2008.

[5] A. Abrardo, D. Dardari, M. DI Renzo, and X. Qian, "MIMO Interference Channels Assisted by Reconfigurable Intelligent Surfaces: Mutual Coupling Aware Sum-Rate Optimization Based on a Mutual Impedance Channel Model," *IEEE Wireless Communications Letters*, vol. 10, no. 12, pp. 2624–2628, 2021.

[6] M. Di Renzo, A. Zappone, M. Debbah, M.-S. Alouini, C. Yuen, J. De Rosny, and S. Tretyakov, "Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead," *IEEE journal on selected areas in communications*, vol. 38, no. 11, pp. 2450–2525, 2020.

[7] A. Albanese, G. Encinas-Lago, V. Sciancalepore, X. Costa-Pérez, D.-T. Phan-Huy, and S. Ros, "RIS-aware indoor network planning: The rennes railway station case," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 2028–2034.

[8] G. Encinas-Lago, A. Albanese, V. Sciancalepore, M. Di Renzo, and X. Costa-Pérez, "Unlocking metasurface practicality for B5G networks: AI-assisted RIS planning," in *GLOBECOM 2023-2023 IEEE Global Communications Conference*. IEEE, 2023, pp. 6560–6566.

[9] X. Gan, C. Huang, Z. Yang, X. Chen, J. He, Z. Zhang, C. Yuen, Y. L. Guan, and M. Debbah, "Coverage and rate analysis for integrated sensing and communication networks," *IEEE Journal on Selected Areas in Communications*, 2024.

[10] J. An, H. Li, D. W. K. Ng, and C. Yuen, "Fundamental detection probability vs. achievable rate tradeoff in integrated sensing and communication systems," *IEEE Transactions on Wireless Communications*, vol. 22, no. 12, pp. 9835–9853, 2023.

[11] A. Albanese, V. Sciancalepore, A. Banchs, and X. Costa-Perez, "LOKO: Localization-aware roll-out planning for future mobile networks," *IEEE Transactions on Mobile Computing*, pp. 1–1, 2022.

[12] T. Ma, Y. Xiao, X. Lei, L. Zhang, Y. Niu, and G. K. Karagiannidis, "Reconfigurable intelligent surface-assisted localization: Technologies, challenges, and the road ahead," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 1430–1451, 2023.

[13] R. Liu, M. Li, H. Luo, Q. Liu, and A. L. Swindlehurst, "Integrated sensing and communication with reconfigurable intelligent surfaces: Opportunities, applications, and future directions," *IEEE Wireless Communications*, vol. 30, no. 1, pp. 50–57, 2023.

[14] Y. Chen, M. Wen, E. Basar, Y.-C. Wu, L. Wang, and W. Liu, "Exploiting reconfigurable intelligent surfaces in edge caching: Joint hybrid beamforming and content placement optimization," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 7799–7812, 2021.

[15] S. Zeng, H. Zhang, B. Di, Z. Han, and L. Song, "Reconfigurable intelligent surface (RIS) assisted wireless coverage extension: RIS orientation and location optimization," *IEEE Communications Letters*, vol. 25, no. 1, pp. 269–273, 2021.

[16] K. Ntontin, A.-A. A. Boulogeorgos, D. G. Selimis, F. I. Lazarakis, A. Alexiou, and S. Chatzinotas, "Reconfigurable intelligent surface optimal placement in millimeter-wave networks," *IEEE Open Journal of the Communications Society*, vol. 2, pp. 704–718, 2021.

[17] S. Hou, N. U. Saqib, S. H. Chae, Z. Dou, and S.-W. Jeon, "Optimal placement of reconfigurable intelligent surface for millimeter-wave indoor communication," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 698–700.

[18] K. Ntontin, D. Selimis, A.-A. A. Boulogeorgos, A. Alexandridis, A. Tsolis, V. Vlachodimitropoulos, and F. Lazarakis, "Optimal reconfigurable intelligent surface placement in millimeter-wave communications," in *2021 15th European Conference on Antennas and Propagation (EuCAP)*, 2021, pp. 1–5.

[19] G. Stratidakis, S. Droulias, and A. Alexiou, "Optimal position and orientation study of reconfigurable intelligent surfaces in a mobile user environment," *IEEE Transactions on Antennas and Propagation*, vol. 70, no. 10, pp. 8863–8871, 2022.

[20] B. P. L. Lau, N. Wijerathne, B. K. K. Ng, and C. Yuen, "Sensor fusion for public space utilization monitoring in a smart city," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 473–481, 2017.

[21] Z. Wei, F. Zhang, S. Chang, Y. Liu, H. Wu, and Z. Feng, "Mmwave radar and vision fusion for object detection in autonomous driving: A review," *Sensors*, vol. 22, no. 7, p. 2542, 2022.

[22] X. Wang, L. Gao, S. Mao, and S. Pandey, "Deepfi: Deep learning for indoor fingerprinting using channel state information," in *2015 IEEE wireless communications and networking conference (WCNC)*. IEEE, 2015, pp. 1666–1671.

[23] C.-K. Wen, L.-S. Tsai, A. Shojaeifard, P.-K. Liao, K.-K. Wong, and C.-B. Chae, "Shaping a smarter electromagnetic landscape: IAB, NCR, and RIS in 5G standard and future 6G," *IEEE Communications Standards Magazine*, vol. 8, no. 1, pp. 72–78, 2024.

[24] Y. Zhang, J. Zhang, M. Di Renzo, H. Xiao, and B. Ai, "Performance analysis of RIS-aided systems with practical phase shift and amplitude response," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 5, pp. 4501–4511, 2021.

[25] G. R. Curry, *Radar System Performance Modeling*. Artech House, 2005.

[26] M. Lichtenstein and T. Young, "The resolution of closely spaced signals," *IEEE Transactions on Information Theory*, vol. 14, no. 2, pp. 288–293, 1968.

[27] J. He, H. Wymeersch, T. Sanguanpuak, O. Silvén, and M. Juntti, "Adaptive beamforming design for mmwave RIS-aided joint localization and communication," in *2020 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*. IEEE, 2020, pp. 1–6.

[28] K. Keykhosravi, G. Seco-Granados, G. C. Alexandropoulos, and H. Wymeersch, "RIS-enabled self-localization: Leveraging controllable reflections with zero access points," in *ICC 2022-IEEE International Conference on Communications*. IEEE, 2022, pp. 2852–2857.

[29] Y. Xing, T. S. Rappaport, and A. Ghosh, "Millimeter wave and sub-thz indoor radio propagation channel measurements, models, and comparisons in an office environment," *IEEE Communications Letters*, vol. 25, no. 10, pp. 3151–3155, 2021.

[30] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. MacCartney, T. S. Rappaport, and A. Alsanie, "Radio propagation path loss models for 5G cellular networks in the 28 GHz and 38 GHz millimeter-wave bands," *IEEE communications magazine*, vol. 52, no. 9, pp. 78–86, 2014.

[31] B. Sihlbom, M. I. Poulakis, and M. Di Renzo, "Reconfigurable intelligent surfaces: Performance assessment through a system-level simulator," *IEEE Wireless Communications*, vol. 30, no. 4, pp. 98–106, 2022.

[32] N.-T. Nguyen and B.-H. Liu, "The mobile sensor deployment problem and the target coverage problem in mobile wireless sensor networks are np-hard," *IEEE Systems Journal*, vol. 13, no. 2, pp. 1312–1315, 2018.

[33] G. Encinas-Lago, A. Albanese, V. Sciancalepore, X. Costa-Pérez, A. Banchs, and D.-T. Phan-Huy, "A cost-effective riss deployment to abate the coverage problem in b5g networks," *IEEE Transactions on Wireless Communications*, 2024.

[34] C. J. C. H. Watkins, "Learning from delayed rewards," 1989.

[35] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, pp. 279–292, 1992.

[36] F. Zhao and L. J. Guibas, *Information processing in sensor networks: Second international workshop, IPSN 2003: Proceedings*. Springer, 2003.

[37] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position error bound for UWB localization in dense cluttered environments," *IEEE transactions on aerospace and electronic systems*, vol. 44, no. 2, pp. 613–628, 2008.

[38] X. Shao, L. Cheng, X. Chen, C. Huang, and D. W. K. Ng, "Reconfigurable intelligent surface-aided 6G massive access: Coupled tensor modeling and sparse bayesian learning," *IEEE Transactions on Wireless Communications*, vol. 21, no. 12, pp. 10 145–10 161, 2022.

[39] *Technical Specification Group Radio Access Network; Base Station (BS) radio transmission and reception*, European Telecommunications Standards Institute, 09 2021, 3GPP TS 38.104 version 15.14.0.

[40] Y. Liu, X. Shi, S. He, and Z. Shi, "Prospective positioning architecture and technologies in 5G networks," *IEEE Network*, vol. 31, no. 6, pp. 115–121, 2017.

BIOGRAPHIES

Guillermo Encinas-Lago (M'21) received his M.Sc. degrees in Applied Physics and in Industry 4.0 respectively from Universidad Autónoma de Madrid in 2013 and Univesidad Carlos III de Madrid in 2021, both in Spain, and a Ph.D. in Telecommunications in 2024 from Université Paris-Saclay, France, while employed at NEC Laboratories Europe in Heidelberg, Germany, under the MSCA program. He works now as Senior Researcher in i2Cat, focused on RISs, machine learning, ISAC, and prototyping.

Vincenzo Sciancalepore (M'15–SM'19) received his M.Sc. degree in Telecommunications Engineering and Telematics Engineering in 2011 and 2012. In 2015, he received a double Ph.D. degree. He is a Principal Researcher at NEC Laboratories Europe, focusing his activity on 6G RAN technologies, such as RISs. He is an editor of IEEE TCOM and IEEE TWC.

Henk Wymeersch (S'01, M'05, SM'19, F'24) obtained the Ph.D. degree in Electrical Engineering/Applied Sciences in 2005 from Ghent University, Belgium. He is currently a Professor of Communication Systems with the Department of Electrical Engineering at Chalmers University of Technology, Sweden. Prior to joining Chalmers, he was a postdoctoral researcher from 2005 until 2009 with the Laboratory for Information and Decision Systems at the Massachusetts Institute of Technology. Prof. Wymeersch served as Associate Editor for IEEE Communication Letters, IEEE Transactions on Wireless Communications, and IEEE Transactions on Communications and is currently Senior Member of the IEEE Signal Processing Magazine Editorial Board. During 2019-2021, he was an IEEE Distinguished Lecturer with the Vehicular Technology Society. His current research interests include the convergence of communication and sensing, in a 5G and Beyond 5G context.

Marco Di Renzo (Fellow, IEEE) received the Ph.D. degree in electrical engineering from the University of L'Aquila, Italy, in 2007. Currently, he is a CNRS Research Director (Professor) and the Head of the Intelligent Physical Communications group with the Laboratory of Signals and Systems (L2S) at CNRS & CentraleSupélec, Paris-Saclay University, Paris, France, as well as a Chair Professor in Telecommunications Engineering with the Centre for Telecommunications Research – Department of Engineering, King's College London, London, United Kingdom. He served as the Editor-in-Chief of IEEE Communications Letters, and he currently serves as the Director of Journals of the IEEE Communications Society.

Xavier Costa-Pérez (M'06–SM'18) is ICREA Research Professor, Scientific Director at the i2Cat Research Center and Head of 5G/6G Networks R&D at NEC Laboratories Europe. He has served on the Organizing Committees of several conferences, published papers of high impact and holds tenths of patents. Xavier received his Ph.D. degree in Telecommunications from the Polytechnic University of Catalonia.