

LOKO: Localization-aware Roll-out Planning for Future Mobile Networks

Antonio Albanese, *Student Member, IEEE*, Vincenzo Sciancalepore, *Senior Member, IEEE*,
Albert Banchs, *Senior Member, IEEE*, Xavier Costa-Pérez, *Senior Member, IEEE*

Abstract—The roll-out phase of the next generation of mobile networks (5G) has started and operators are required to devise deployment solutions while pursuing *localization* accuracy maximization. Enabling *location-based services* is expected to be a unique selling point for service providers now able to deliver critical mobile services, e.g., autonomous driving, public safety, remote operations. In this paper, we propose a novel roll-out base station placement solution that, given a *Throughput-Positioning Ratio* (TPR) target, selects the location of new-generation base stations (among *available candidate sites*) such that the *throughput* and *localization accuracy* are jointly maximized. Moving away from the canonical position error bound (PEB) analysis, we develop a realistic framework in which each positioning measurement is affected by errors depending upon the actual wireless channel between the measuring base station and the target device. Our solution, referred to as LOKO, is a fast-converging algorithm that can be readily applied to current 5G (or future) roll-out processes. LOKO is validated by means of an exhaustive simulation campaign considering *real existing deployments* of a major European network operator as well as synthetic scenarios.

Index Terms—Optimization, Semidefinite Relaxation (SDR), Roll-out, Planning, Localization, B5G, 5G.



1 INTRODUCTION

The next generation of cellular networks (5G) is being bootstrapped by major network operators. Mobile operators are promptly responding to the ever-increasing demand of ubiquitous and reliable cellular communications but facing severe challenges in nurturing new business initiatives with innovative services [1]. Constructing a new set of user contextual information might be the winning strategy to kick-start novel business opportunities: *accurate network-based positioning* is the epitome of such advanced features. Thus, a novel, agile and efficient roll-out strategy is mandatory to support the unprecedentedly high requirements the next generation of mobile networks is expected to deliver. However, positioning accuracy may come at an unaffordable implementation cost, hence operators need to find a viable solution to keep the Total Cost of Ownership (TCO) of the next-generation Radio Access Network (RAN) as low as possible while minimizing the impact on the existing infrastructure [2].

Localization accuracy is one of the widely accepted Key Performance Indicators (KPIs) for 5G services and market technology enablers. New common commercial use cases require a (horizontal) positioning error within 3m or 10m in indoor or outdoor scenarios, respectively, for at least 80% of the User Equipments (UEs) in the service area [3]. In addition to those general use cases, the option for *higher accuracy positioning* was defined in [4] to guarantee UEs localization with low latency and high reliability. For instance, 5G should be able to retrieve the position of vehicles moving at a speed of (up to) 200 km/h with less than 1 m accuracy

in more than 95 % of service area. Finer positioning capabilities are an added-value that would benefit forthcoming network-based use-cases, e.g., tracking of goods and products within kitting processes in manufacturing, support for automated guided vehicles (AGVs) in logistic operations or victims and people in distress localization during search-and-rescue operations [5], [6].

Although Global Navigation Satellite Systems (GNSSs) may provide reasonably accurate users positions, their navigation signals may incur in severe attenuation originated by e.g., debris, tree foliage, metal or wooden objects or shadowing in urban canyons thereby calling for a more reliable backup solution such as an enhanced network-based localization system. Furthermore, GNSSs are managed by third parties out of the scope of network operators, which seek increasingly higher control on such information to pursue the above-mentioned and all the upcoming business opportunities. In particular, 3rd Generation Partnership Project (3GPP) standard guidelines do not define any Application Programming Interface (API) for the purpose, thus requiring users who would willingly share their GNSS-based information to keep a client application running on their UEs, which would have a non-negligible impact on their battery life. By boosting network-based localization, network operators would benefit from the possibility of retrieving users' positions with unprecedented high accuracy without relying on UEs sharing their GNSS-based location on the data plane, therefore opening up a series of new business revenue streams for the heavily location-oriented services of the future network generations [7].

The current fifth generation (5G) cellular network introduces advanced radio features, like ultra-wide bandwidth for better time resolution, additional frequency bands and massive Multi-Input-Multi-Output (MIMO) with millimeter-wave (mmWave) communications for Angle-of-Arrival (AoA) estimation, propelling a new wave of networks able to directly infer positioning information with high accuracy without making use of third-party technologies [8]. Indeed, 5G is expected to provide enhanced po-

A. Albanese is with NEC Laboratories Europe and University Carlos III of Madrid, 28911 Leganés, Spain.

V. Sciancalepore is with NEC Laboratories Europe, 69115 Heidelberg, Germany.

Xavier Costa-Pérez is with i2cat, NEC Laboratories Europe and ICREA, 08034 Barcelona, Spain.

Albert Banchs is with University Carlos III of Madrid and IMDEA Networks Institute, 28911 Leganés, Spain.

Email of corresponding author: antonio.albanese@neclab.eu.

sitioning capabilities, achieving (horizontal) accuracy up to 0.3m (99.9% availability) with 15ms latency after its complete roll-out [9]. From an architectural perspective, multiple radio access technologies (RATs) can be used to deliver diverse localization performances based on given requirements whereas the separation of user and control planes enables real-time positioning in the user plane by virtue of the ever-growing edge cloud [10]. Furthermore, future beyond-5G (B5G) networks will benefit from the ground-breaking Reconfigurable Intelligent Surface (RIS) technology, namely passive reflectarrays able to backscatter and phase-shift the electromagnetic waves impinging upon them by individually controlling their antenna elements [11]. By doing so, RISs transform the channel into an optimization variable overcoming the traditional postulate of an uncontrollable propagation environment. Although modern localization techniques can also take advantage of measurements from non-line-of-sight (NLoS) links, e.g., through simultaneous localization and sensing (SLAM), RISs will help circumvent line-of-sight (LoS) blockage, thus drastically improving the overall localization accuracy even in harsh propagation environments [12].

Even if such new features bring in major accuracy improvements on a single-link basis per se (e.g., leveraging measurements provided by a single Base Station (BS)), the resulting performance still depends on the base stations locations. Thus, positioning performance needs to be thoroughly addressed during the network planning phase to unleash its full potential when leveraging on measurements performed by multiple BSs. Therefore, a proper RAN deployment process becomes an impelling task: while previous RAN design solutions in the literature focus only on throughput maximization and coverage optimization, they mostly disrupt positioning accuracy performance due to the geometric dispersion of base stations or unexpected channel conditions [13]. Furthermore, with the advent of network slicing, network operators may shift their focus from the classical enhanced Mobile BroadBand (eMBB) use case, whose performance is mainly related to network throughput, depending on their major customers needs [14]. For instance, a tenant from the automotive vertical sector may have lower throughput requirements in favor of higher vehicle positioning accuracy, especially in the context of the rising autonomous driving trend. In this regard, operators may take advantage of 5G (and beyond) base stations to increase the accuracy of the positioning procedures while still keeping resource allocation efficiency and end-user rates above reasonable quality levels.

In this paper, we propose a pioneering method to roll out a mobile network selecting the best base station locations among available candidate sites defined by the network operator in a pre-negotiation phase with municipal and private parties, called Localization-aware radio planning Keying on throughput Optimization (LOKO). Our solution is the first-of-its-kind to jointly pursue unprecedented Key Performance Indicators (KPIs), namely user positioning, throughput and fairness while considering existing network deployment limitations. In particular, differently from the canonical approach in the literature, we do not neglect the dependence of single-link positioning accuracy on the actual distance between the measuring BS and the target User Equipment (UE), thus ruling out most trivial deployment solutions, e.g. regular polygon-shaped deployments. The contributions of our paper can be summarized as follows: we *i*) first bring the localization performance into the picture of new RAN planning, *ii*) derive a suitable formulation of the *Position Error Bound (PEB)* for Time-

of-Arrival (ToA) measurements accounting for the dependence of each error on the actual related measurement, *iii*) cast the roll-out process of upcoming network generation (e.g., 5G) into a novel optimization problem that jointly accounts for PEB optimization and throughput maximization, *iv*) introduce a novel parameter, dubbed as *Throughput-Positioning Ratio (TPR)*, which trades off between such KPIs according to the operator requirements, *v*) devise a Block Coordinate-based algorithm that iteratively applies the Semidefinite Relaxation and Bisection techniques to different routines, *vi*) provide convergence proof of the proposed algorithm and *vii*) carry out an exhaustive simulation campaign based on synthetic and real base station deployments (from a major European network operator) to prove the effectiveness of our approach.

The remainder of this paper is as follows. Section 2 provides a summary of 5G-NR deployment modes and a primer on cellular-based localization; Section 3 introduces the LOKO framework along with the novel concept of Throughput-Positioning Ratio (TPR), defining the desired flavour of the joint BS deployment. Section 4 presents the radio planning problem and proceeds by splitting it in two constituent subproblems, while Section 5 describes their algorithmic solution, which is in turn numerically validated by simulations presented in Section 6. Section 7 includes a survey of the related literature and, finally, Section 8 concludes the paper.

NOTATION

We let \mathbb{R} denote the set of real numbers. We use \mathbb{R}_+ , \mathbb{R}^n , and $\mathbb{R}^{n \times m}$ to represent the sets of non-negative real numbers, n -dimensional real vectors, and $m \times n$ real matrices, respectively. Vectors are denoted by default as column vectors and written in bold font. Notable examples are $\mathbf{1}_n$ and $\mathbf{1}_{n \times n}$, which respectively denote the unit vector and the unit square matrix of size n , while \mathbf{I}_n is the identity matrix in $\mathbb{R}^{n \times n}$. Subscripts represent an element in a vector and superscripts elements in a sequence. For instance, $\mathbf{x}^{(m)} = [x_1^{(t)}, \dots, x_n^{(t)}]^\top$ is a vector from \mathbb{R}^n and $x_i^{(m)}$ is its i th component. Operator $(\cdot)^\top$ represents the transpose operation, $\text{Tr}(\cdot)$ denotes the trace of a square matrix, $\|\cdot\|$ indicates the L2-norm of a vector and ∇ denotes the gradient operator. Finally, \odot indicates the Hadamard or element-wise product between two matrices while $\text{diag}(\cdot)$ returns the main diagonal of its argument.

2 PRELIMINARIES

Hereafter, we detail the state-of-the-art of legacy and 5G RAN deployments thus shedding light on recent positioning findings.

2.1 Non-Standalone Deployment: Legacy and 5G RANs

Recently, 3GPP has defined the Non-Standalone (NSA) deployment mode as a 5G-NR deployment option that relies on the existing LTE core network (EPC) while applying 5G-NR features to user plane functions only [15]. More and more operators prefer it in the short-term: they may ease the complexity of deploying a nationwide Standalone (SA) 5G network from scratch (as opposed to NSA) by performing only a selective deployment of 5G-NR base stations (gNBs) with no need of a proper 5G Core (5GC). Therefore, we consider hybrid scenarios with legacy¹ base stations

1. Our main findings may be easily applied to 5G hybrid scenarios employing any legacy radio access technology, including LTE, 3G and so on. For the sake of simplicity, in this paper we only model the LTE technology.

already in place. Please note that our solution may be readily applied to SA deployment mode, as 5G-only networks are a subset of such hybrid scenarios.

Hybrid Deployment. To guarantee backwards compatibility in the new RAN design, low frequencies—employed by LTE—will continue being used to provide sufficient signal level in rural, suburban and urban areas [16], while most network operators have favored the higher C-Band for their first 5G-NR deployments with up to five-time channel bandwidth increase compared to LTE [17]. Indeed, a higher bandwidth associated with a higher spectral efficiency allows gNBs to offer much larger throughput w.r.t. a classical eNB. Additionally, a higher bandwidth allows for a proportionally higher resolution in the delay domain, making multipath easier to address, thus alleviating the measurement error in range-based positioning, see e.g. the wide-spread usage of Ultra Wide Bandwidth (UWB) signals for localization [18].

2.2 Time-based positioning

Time-based positioning is a fundamental positioning technique, which combines Time-of-Flight (ToF) measurements related to a set of (at least three) base stations to derive the location of some User Equipment (UE) by means of the multilateration technique. To this aim, the UE receives periodic Positioning Reference Signals (PRSs) from the serving and the adjacent base stations and records their respective Time of Arrivals (ToAs)². PRSs are sent within Positioning subframes, which are designed as Low Interference Subframes (LISs), i.e. with no concurrent data transmission over the Physical Downlink Shared CHannel (PDSCH). Therefore, in perfectly synchronized networks, PRSs are isolated from data transmission and are only exposed to the interference generated by other cell PRSs (with the same pattern index). Dedicating time and bandwidth resources (in terms of Physical Resource Blocks (PRBs)) to PRSs has an impact on the achievable network throughput, which is usually lower than 1% in normal operating conditions, as by 3GPP recommendations [21]. Therefore, in our analysis, we neglect such effect as it would tie the derivation to the specific implementation of the multilateration technique³. Below we derive the lower bound on the positioning accuracy provided by a technique of this kind, which in the rest of the paper we employ as positioning performance metric.

Time-of-Arrival measurements. Let us consider a cellular network consisting of B base stations (BSs) deployed at coordinates $\mathbf{q}_i = (x_i, y_i)^\top$ and one UE performing ranging measurements⁴ from all available BSs in order to infer its own

2. For the sake of tractability, we focus our analysis on ToA only. It is worth pointing out that Time-Difference-of-Arrival (TDoA) eliminates the need for knowing the absolute transmission times of PRSs by subtracting the ToAs of a reference base station from the collection of measurements. However, ToA and TDoA provide the same theoretical lower bound in conventional scenarios. Besides, the PEB in NLoS conditions with some a-priori knowledge on the biases is lower-bounded by the PEB attained in perfect LoS conditions. We refer the reader to [19], [20] for further details.

3. Nonetheless, our problem formulation may be extended to account for PRSs overhead by suitably scaling down the achievable throughput expression in the problem formulation.

4. We recognize that in practice several options are available to improve the performance of a localization technique solely based on time measurements, e.g. MIMO systems for estimating Angle-of-Arrival (AoA) and Angle-of-Departure (AoD), cooperation among base stations, and so forth. However, in this paper we propose the first localization-aware deployment solution and show the implications of such aspects on real deployment snapshots. Therefore, we argue that the inclusion of the above-mentioned positioning techniques is out of scope.

geographical coordinates $\mathbf{p} = (x, y)^\top$. Let us denote the set of ToA range measurements as:

$$\tilde{\varrho}_i(\mathbf{p}) = d_i(\mathbf{p}) + l_i + n_i, \quad \forall i \in \mathcal{B}, \quad (1)$$

where $d_i(\mathbf{p}) = \|\mathbf{p} - \mathbf{q}_i\|$ is the true Euclidean distance between the UE and i th base station, l_i is a positive bias modeling the excess path length due to Non-Line-of-Sight (NLoS) propagation conditions, and n_i is a random error.

Following [22], we assume that biases l_i are uniformly distributed between 0 and some λ_i ⁵. This assumption strives to minimize the burden on network operators performing an extensive measurement campaign. Indeed, estimating the statistics of the ToA measurement biases in the area for every possible BSs deployment is a particularly demanding task, which, in turn, leads to a difficult-to-manage PEB expression in the light of its optimization. Therefore, we limit the measurement effort to the maximum bias values and assume their statistics to be uniform between 0 and such values, i.e. λ_i 's. Measurement errors n_i are modeled as independent Gaussian random noises with zero mean and variance σ_i^2 depending on the distance d_i [23], namely $\sigma_i^2 = \sigma_0^2 (d_i/d_0)^{\alpha_i}$, where $\alpha_i \geq 0$ is the path-loss exponent and σ_0^2 is the variance at a reference distance d_0 . Indeed, the higher the distance, the higher the number of indirect propagation paths, thereby exacerbating the task of correctly estimating the time of arrival of the dominant path at the UE side. Moreover, this preserves the generality of our model as it includes the distance-invariant one by setting $\alpha_i = 0$.

The probability density function (pdf) of unbiased range measurements ϱ_i can be written as

$$f_{\varrho_i}(\varrho_i|\mathbf{p}) = \frac{1}{\lambda_i} \left[Q\left(\frac{\varrho_i - d_i(\mathbf{p}) - \lambda_i}{\sigma_i}\right) - Q\left(\frac{\varrho_i - d_i(\mathbf{p})}{\sigma_i}\right) \right], \quad (2)$$

where $Q(z) = \frac{1}{\sqrt{2\pi}} \int_z^{+\infty} e^{-t^2} dt$ is the Gaussian Q-function⁶.

Position Error Bound. To assess the localization performance of cellular networks, we leverage on the *Cramér-Rao Lower Bound* (CRLB), which provides a lower bound on the covariance matrix of any unbiased estimator. This makes our problem completely general, without tailoring it to a specific localization technique (e.g., [24]), while still providing useful insights on the expected performance of a positioning system deployed in the field [25]. The choice of assessing localization performances by means of the PEB stems from the need for a general and reasonably compact expression to optimize the localization capabilities of the network, which is independent of the specific implemented localization solution. This approach goes along the lines of making use of the capacity definition in place of the sum-rate, which in reality would be affected by, e.g., the specific scheduling and resource allocation at the BSs. Let us define the estimated UE position as $\hat{\mathbf{p}} = (\hat{x}, \hat{y})^\top$ based on range measurements vector $\boldsymbol{\varrho} = (\varrho_1, \dots, \varrho_B)^\top$. The covariance matrix of $\hat{\mathbf{p}}$ satisfies:

$$\text{cov}(\boldsymbol{\varrho}) = \mathbb{E}_{\boldsymbol{\varrho}} [(\mathbf{p} - \hat{\mathbf{p}})(\mathbf{p} - \hat{\mathbf{p}})^\top] \succeq \mathbf{J}_p^{-1}, \quad (3)$$

where $\mathbb{E}_{\boldsymbol{\varrho}}[\cdot]$ represents the expectation over $\boldsymbol{\varrho}$, \succeq indicates that $\text{cov}(\boldsymbol{\varrho}) - \mathbf{J}_p^{-1}$ is positive semidefinite and \mathbf{J}_p is the Fisher Infor-

5. Typically, the only a-priori information about the environment is the maximum bias λ_i while a detailed description of its distribution may not be available.

6. A closed-form expression for the TDoA case is not available. However, numerical results show that the lower the lumped error $l + n$ on the reference ToA compared to the lumped error on the current ToA, the better Eq. (2) approximates the pdf of the resulting TDoA measurements.

mation Matrix (FIM) for a given \mathbf{p} and base station deployment $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_B)^\top$, defined as:

$$\mathbf{J}_p = \mathbb{E}_{\boldsymbol{\rho}} [(\nabla_{\mathbf{p}} \ln f(\boldsymbol{\rho}|\mathbf{p})) (\nabla_{\mathbf{p}} \ln f(\boldsymbol{\rho}|\mathbf{p}))^\top]. \quad (4)$$

Note that, since range measurements are assumed to be independent, we have $f(\boldsymbol{\rho}|\mathbf{p}) = \prod_{i=1}^B f_i(\rho_i|\mathbf{p})$. From Eq. (3), it yields the following

$$\sqrt{\mathbb{E}_{\boldsymbol{\rho}} [(x - \hat{x})^2 + (y - \hat{y})^2]} \geq \sqrt{\text{Tr}(\mathbf{J}_p^{-1})}. \quad (5)$$

The value $\sqrt{\text{Tr}(\mathbf{J}_p^{-1})}$ is dubbed as *Position Error Bound* (PEB) and represents a lower bound on the mean-square error (MSE) of the distance between the UE position and its corresponding estimate. The PEB of ToA measurements with uniformly distributed measurement biases can be written as the following ([22])⁷

$$\beta(\mathbf{p}) = \left(\frac{\sum_{i \in \mathcal{B}} \nu_i}{\sum_{\substack{i < j, \\ i, j \in \mathcal{B}}} \nu_i \nu_j \sin^2(\theta_j - \theta_i)} \right)^{\frac{1}{2}}, \quad (6)$$

where ν_i is a function of d_i and λ_i :

$$\nu_i(d_i, \lambda_i) = \frac{1}{\lambda_i \sigma_i(d_i) \pi \sqrt{2}} \int_{-\infty}^{+\infty} h(y, \lambda_i, d_i) dy, \quad (7)$$

with $d_i = d_i(\mathbf{p})$, $\lambda_i = \lambda_i(\mathbf{p})$, $\theta_i = \theta_i(\mathbf{p})$ the angle between \mathbf{p} and \mathbf{q}_i w.r.t. the horizontal axis, and $h(y, \lambda, d) =$

$$\frac{\left(e^{-\left(y + \frac{\lambda}{\sigma(d)\sqrt{2}}\right)^2} \left(1 + \frac{\alpha\lambda}{2d} + \frac{\alpha\sigma(d)}{d\sqrt{(2)}} y \right) - e^{-y^2} \left(1 + \frac{\alpha\sigma(d)}{d\sqrt{2}} y \right) \right)^2}{Q\left(\sqrt{2}y\right) - Q\left(\sqrt{2}y + \frac{\lambda}{\sigma(d)}\right)}. \quad (8)$$

We would like to point out that accounting for the dependence of the measurement errors on the actual distance is key for making the PEB a reasonable metric for our deployment solution. Indeed, under the assumption of independent errors, namely for $\alpha_i = 0$, $\forall i \in \mathcal{B}$, it is well-known that the lowest value of PEB occurs at the center of a regular B -sided polygon of any side length [26]. Therefore, one trivial solution to the PEB optimization problem would be to spread the BSs regularly around the deployment area far enough from each UE to allow approximating their position with the polygon center. On the contrary, if single-link errors depend on the actual distance, this kind of solution is ruled out and some other deployment schemes not necessarily regular around each UE might be possible, as shown in the following.

The PEB in Eq. (6) depends on a twofold factor: *i*) the propagation conditions experienced by the transmitted signals through weights ν_i and *ii*) the angular positions of each base station (BS) w.r.t. the UE. In realistic deployments, the performance gain expected from 5G-NR with respect to the previous generation is difficult to achieve only by resorting on physical layer improvements but rather requires a paradigm shift in the RAN roll-out by considering the positioning performance in Eq. (6) as a target KPI in the planning phase.

7. This definition is not limited to UWB, but it applies to any ToA-based localization system regardless of its transmit power, operating bandwidth and propagation environment.

2.3 Measurement impairments and randomness

As described in Section 2.2, NLoS conditions introduce a bias in the range measurements due to the blockage of the Line-of-Sight (LoS) direct path with the serving base station, e.g. in presence of shadowing. In such case, the UE measures a positively biased distance since the dominant received signal propagates along an indirect path, which is longer than the obstructed direct one. Travelling a longer path, the signal experiences a higher attenuation that is usually modeled according to a log-normal distribution with σ_s standard deviation.

It is worth pointing out that the direct path may still be received by the UE if the additional attenuation introduced by the blocking obstacles is not enough to lower the received power below the UE sensitivity. Nevertheless, the refraction experienced by such path propagating through the obstacles leads to an additional delay, whose value depends on the refractive index of the obstacles material. Therefore, there is a close proportional link between the bias distribution in the range measurement process and the magnitude of the wireless channel shadowing, namely between λ and σ_s , which we take into account in our numerical evaluation in Section 6.

3 LOKO FRAMEWORK OVERVIEW

A proper radio planning lies in selecting BSs locations and configuration parameters in order to provide full coverage and meet the required Quality of Service (QoS) in a target area, while keeping the network operator CAPEX and OPEX as low as possible. State-of-the-art solutions including commercial radio planning software packages provide heuristic solutions of the problem in its classical flavour, which takes no account of users' localization. With LOKO, we take a step forward by designing an adaptive technique to jointly optimize network throughput, user fairness and positioning performances.

Network integration. Operators willing to incorporate 5G-NR in their live network likely have got an up-and-running legacy deployment, such as LTE or 3G. The most convenient way to bootstrap 5G-NR deployment is to employ the NSA option, as discussed in Section 2.1. Therefore, LOKO takes the pre-existing LTE network into account and improves upon its throughput and positioning performances, which represent the guaranteed baseline.

Deployment settings. In a realistic operation scenario, network operators must not exceed their budget, namely they can afford only a limited given number of new base stations, while meeting logistical and administrative constraints during the deployment phase. On the one hand, they are required to comply with municipal restrictions in terms of electromagnetic emissions and public land occupation. On the other hand, they seek one-to-one agreements with private parties to install network equipment in their facilities. In addition to the administrative regulations, there are also physical constraints to deal with. For instance, when deploying a base station on the roof of a building, there might be only a specific location on the roof where network operators could physically place the base station. Likewise, not all buildings in a city may allow a base station on their roof. Thus, LOKO takes as input the set of candidate sites that satisfy the above-mentioned requirements, and selects the corresponding subset wherein new base stations will be deployed without violating the budget constraint. We would like to underline that there is no intrinsic limitation on the set of candidate sites, which could

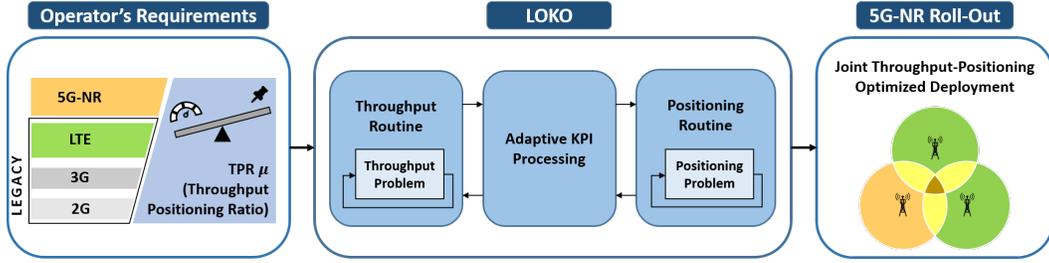


Fig. 1: Overview of LOKO's building blocks

TABLE 1: Terminology

Acronym	Parameter	Acronym	Parameter
RAT	Radio Access Technology	RAN	Radio Access Network
ToA	Time-of-Arrival	GNSS	Global Navigation Satellite System
CRLB	Cramér-Rao Lower Bound	PEB	Position Error Bound
BS	Base Station	eNB/gNB	LTE/5G Base Station
UE	User Equipment	CS	Candidate site
TPR	Throughput-Positioning Ratio	SDR	Semidefinite Relaxation
BCA/D	Block Coordinate Ascend/Descend	FP	Fractional Programming

possibly include any location in the service area, e.g. all buildings in a city. Therefore, considering a set of candidate sites does not introduce any loss of generality in our proposed solution. Note that this set contains also the sites with existing LTE BSs to account for the possibility of co-locating LTE and 5G BSs and thus reducing installation costs; however, there is no guarantee that the set of existing LTE locations would satisfy the positioning accuracy requirements of the operator in the 5G era. We propose to sample the target area by means of T test points, wherein a typical UE should meet the desired requirements in terms of throughput and positioning performances. The test point distribution should match the expected distribution of cellular users in such area. Nevertheless, our solution is not tailored to a specific set of test points and can be readily applied to any user distribution.

Flavour tuning. Throughput and localization performances are different metrics that employ different units of measurements. As there is a trade-off between throughput and location accuracy, this raises the need for a tuning parameter μ , dubbed as Throughput-Positioning Ratio (TPR), which reflects the desired flavour of the planning solution. Indeed, TPR μ takes into account how much an operator values a unit gain towards throughput or localization performances, in other words how many *bit/s* would it put at stake to improve the PEB by 1 meter.

Iterative Solution. As shown in Fig. 1, LOKO takes as input the above-mentioned sites and the desired planning flavour and performs an iterative procedure, which quickly converges to a planning solution fulfilling the QoS requirements of the operator.

Lastly, Table 1 lists the terminology used throughout the manuscript for better readability.

4 SYSTEM MODEL

We formalize the overall problem in order to strike the perfect balance between throughput and localization performances experienced by a typical user in the deployment area by means of a tunable parameter, namely Throughput Positioning Ratio (TPR). Specifically, we first formulate it as a joint optimization problem and discuss its tractability. Then, we boil down its complexity by considering two constituent problems and manipulating them to obtain algorithmic sub-optimal solutions, which are described in detail in Sections 5.2 and 5.3 respectively.

4.1 Joint Radio Planning Problem

As previously discussed in Section 3, our target area encompasses pre-deployed legacy base stations (in our example, LTE) providing the performance baseline that we aim to improve by means of our solution, namely LOKO.

Let us characterize the set of base stations \mathcal{B} as the union of two disjoint subsets (or tiers) \mathcal{E} and \mathcal{G} of LTE (eNBs) and 5G-NR (gNBs) base stations, respectively. The former encloses all already deployed eNBs while the latter includes all available gNBs that can be deployed without exceeding the allocated budget by the network provider. Note that we assume full frequency reuse, and same transmission parameters and propagation conditions within a single tier⁸, while considering no interference between the two, as per Section 2.1. Therefore, we respectively denote by P_k , f_k , W_k , α_k and σ_{s_k} the transmit power, central frequency, bandwidth, propagation factor and shadowing standard deviation associated to tier k , where k can exclusively take values e or g according to the corresponding tier \mathcal{E} or \mathcal{G} , respectively. We seek a *hybrid user association strategy*, namely a typical UE associates to the base station showing the highest downlink Signal-to-Interference-plus-Noise-Ratio (SINR) among the available ones, given that it allows satisfying the positioning accuracy requirements⁹.

Let us write the average downlink SINR experienced in the slow-fading regime by a typical UE at coordinates \mathbf{p} associated to base station i belonging to tier k as

$$\gamma_i(\mathbf{p}) = \frac{P_k g_i(\mathbf{p})}{P_k g'_i(\mathbf{p}) + N_0 W_k}, \quad (9)$$

where $g_i(\mathbf{p}) = (4\pi f_k d_i(\mathbf{p})/v)^{-\alpha} \exp(-\sigma_{s_k}^2/2\xi^2)$, with $\xi = 10/\log(10)$ and v indicating the speed of light in vacuum, is the average channel attenuation in presence of log-normal shadowing, N_0 is the two-sided noise power spectral density and $g'_i(\mathbf{p})$ denotes the power-normalized interference experienced by the UE, namely $g'_i(\mathbf{p}) = \sum_{n \in \mathcal{K} \setminus \{i\}} g_n(\mathbf{p})$, where $\mathcal{K} = \{\mathcal{E}, \mathcal{G}\}$ indicates which subset (5G or LTE) base station i belongs to. Note that interference arises only from active base stations belonging to the same tier as the serving one, as previously discussed in Section 2.1. Beamforming capabilities of MIMO systems could reduce the interference between neighboring base stations operating on the same bandwidth. However, especially in dense-urban scenarios, UEs may be spatially spread around each BS, thereby leading to wide beams at the limit corresponding to a Single-Input-Single-Output (SISO) radio configuration.

⁸. These assumptions serve the purpose of a more compact notation, although our analysis may be readily extended.

⁹. As per a conventional maximum-SINR association strategy, the UE prefers to associate to a gNB if another eNB offers the same SINR. When two or more gNBs are eligible for association, the UE performs a random decision.

Decision variables. To facilitate the installation procedure, we assume that base stations may be deployed only at operator-defined candidate sites (CSs), thus calling for coordinate discretisation. It is worth pointing out that LOKO may likewise be executed even if there is not any a-priori knowledge of CSs by taking as input any sampling of the target area. We name \mathcal{S} the set of such CSs with coordinates $\mathbf{s}_j \in \Omega$, $\forall j \in \mathcal{S}$, where $\Omega \in \mathbb{R}^2$ is the target deployment area, being S the cardinality of \mathcal{S} . For example, \mathcal{S} may be a superset of the sites where LTE base stations have already been deployed. Indeed, eNBs coordinates are fed into the problem, and kept fixed during the optimization process.

Our planning solution outputs the set of CSs wherein gNBs will be deployed, while providing the association between the typical UE in each test point and its optimal base station under the assumption of a single UE per test point. Therefore, we introduce decision variables $x_j \in \{0, 1\}$ and $a_{tj} \in \{0, 1\}$ indicating whether a gNB is deployed at CS j and the association between a typical UE at test point t and CS j , respectively.

Network performances. In principle, to calculate the network throughput experienced at each test point, we should average the user throughput over time. However, user rate depends on the serving base station scheduler, which assigns the optimal Modulation and Coding Scheme (MCS) given the current channel quality. Indeed, for the sake of tractability, we only consider Shannon's channel capacity $W_k \log_2(1 + \gamma_i(\mathbf{p}_t))$, assuming that the base station schedulers perform near-optimal MCS assignments and the entire bandwidth of the best station is assigned to the test user. It provides an upper bound on the throughput that a typical UE experiences at test point t with coordinates $\mathbf{p}_t \in \Omega$ while being served by base station i . As far as positioning performances go, we refer to the PEB defined in Section 2.2¹⁰. By using this formulation, we depart from the conventional assumption that each measurement error is independent of the actual measurement thereby accounting for the additional attenuation and fading experienced by UEs that are further away from the measuring BS, thus improving the validity of our system model. Said that a typical UE receives PRSs from active BSs, we need to discriminate only between LTE and 5G network association, relaxing the need of finely optimizing association variables a_{tj} .

Throughput Positioning Ratio (TPR μ). To properly account for heterogeneous objectives, we need to introduce in the objective function a scalar factor μ (TPR) that trades off between throughput and positioning performances. In the extreme case of $\mu = 0$, our optimization will provide only throughput guarantees. Conversely, when $\mu \rightarrow \infty$, the solution will optimize positioning performances regardless of the experienced throughput values. A detailed evaluation of such a parameter is provided in Section 6.2 and Table 5.

Problem formulation. We formulate the joint throughput-positioning planning problem¹¹. It is worth pointing out that our joint goal is to maximize the fairness among test points, both throughput and positioning-wise. Note that this objective maps onto a max-min problem as follows

10. We only refer to Single-Input-Single-Output (SISO) Capacity. Although 5G-NR has MIMO capabilities, we spare the additional model complexity as it would provide analogous results at least with quasi-uniform user distributions.

11. The overall deployment cost or power consumption vary according to the selected sites. Although we do not explicitly account for them, the model can be easily extended by adding two corresponding linear budget constraints, with little to no modifications to the inner loops described in Sections 5.2 and 5.3.

Problem 1 (Joint Throughput-Positioning Radio Planning).

$$\max_{\mathbf{A}, \mathbf{x}} \quad \min_{t \in \mathcal{T}} (\mathbf{r} - \mu \mathbf{b}) \quad (10a)$$

$$\text{subject to} \quad \mathbf{A} \leq \mathbf{1}_T \mathbf{x}^\top, \quad (10b)$$

$$\mathbf{A} \mathbf{1}_S \leq \mathbf{1}_T, \quad (10c)$$

$$\mathbf{x}^\top \mathbf{1}_S \leq G, \quad (10d)$$

$$\mathbf{x} \in \{0, 1\}^S, \quad \mathbf{A} \in \{0, 1\}^{T \times S}, \quad (10e)$$

wherein \mathbf{A} , with $\{\mathbf{A}\}_{tj} = a_{tj}$, is the association decision matrix and $\mathbf{x} = [x_1, \dots, x_S]^\top$ is the deployment decision vector. Constraint (10b) states that a gNB must be deployed at a certain CS as long as one test point is associated to it, while constraint (10c) reflects the fact that each test point can be associated to at most one CS (thus base station), as prescribed by 3GPP. Moreover, constraint (10d) enforces the budget constraint by limiting the maximum number of deployed gNBs to G , where $G = |\mathcal{G}|$ is the cardinality of \mathcal{G} . Objective function (10a) counts two contributions, namely \mathbf{r} and \mathbf{b} , which are shortenings for $\mathbf{r}(\mathbf{A}, \mathbf{x})$ and $\mathbf{b}(\mathbf{A}, \mathbf{x})$, respectively. The former considers the throughput performance and its elements can be written as

$$r_t = \sum_{j \in \mathcal{S}} \left(a_{tj} \log_2 \left(1 + \frac{g_{tj}}{g'_{tj} + N'} \right) \right) W_g + a'_t c_t, \quad (11)$$

where $g_{tj} = g_j(\mathbf{p}_t)$, $g'_{tj} = \sum_{n \in \mathcal{S} \setminus \{j\}} g_{tn} x_n$ indicates the normalized interference experienced by a typical UE at site t given deployment decision \mathbf{x} , $N' = N_0 W_g / P_g$ is the normalized noise power, $a'_t = 1 - \sum_j a_{tj}$, and c_t is the throughput provided at \mathbf{p}_t by the pre-deployed eNBs, namely

$$c_t = \max_{m \in \mathcal{E}} W_e \log_2 (1 + \gamma_m(\mathbf{p}_t)). \quad (12)$$

Note that, whenever a UE is better off connecting to the LTE network, there is no need in our solution to keep track of the specific eNB, as the best association is defined beforehand by the legacy deployment.

The second contribution \mathbf{b} takes into account the positioning performance. Its elements are of the form

$$b_t = (1 - a'_t) \left(\frac{\boldsymbol{\nu}_t^\top \mathbf{x}}{\mathbf{x}^\top \mathbf{F}_t \mathbf{x}} \right) + a'_t u_t^2, \quad (13)$$

where $\boldsymbol{\nu}_t = [\nu_{1t}, \dots, \nu_{St}]^\top$, with $\nu_{jt} = \nu_j(\mathbf{p}_t)$, \mathbf{F}_t is a strictly upper triangular matrix so as $\{\mathbf{F}_t\}_{i,j} = \nu_{it} \nu_{jt} \sin^2(\theta_{jt} - \theta_{it}) \forall i < j$, with $\theta_{jt} = \theta_j(\mathbf{p}_t)$ and u_t is the PEB given by the LTE deployment as per Eq. (6).

Analytical tractability. Even reduced versions of Problem 1 are NP-Hard, thus making it difficult to find its exact solution. For instance, let us consider a scenario where all users experience the same SINR towards all BSs and optimize for throughput only. In this case, the problem is equivalent to a bin packing problem, which is known to be NP-hard. Therefore, we develop a heuristic solution by means of convex relaxation techniques.

Binary constraint (10e) makes Problem 1 combinatorial. Although after relaxing this constraint the set of constraints is convex, the joint objective function is still non-convex, being the sum of non-convex components. Each component calls for a different solution methodology, which makes it impractical to tackle the joint optimization problem directly. In the following, we break down the original problem into two problems and derive a sub-optimal solution for each of them. Such solutions constitute the buildings blocks of LOKO as shown in Fig. 1,

which is eventually able to obtain a solution for the original joint optimization problem.

4.2 Problem separation

3GPP standards have heretofore favored throughput over other performance figures, making it the de-facto most compelling parameter for any cellular network deployment. However, more recently we are witnessing a reversal of the trend with the ongoing standardization of 3GPP fifth generation. In particular, 5G promises to improve several neglected figures such as end-to-end latency, support for massive IoT and enhanced positioning accuracy.

Throughput Problem. Nevertheless, squeezing the most throughput out of a network deployment remains a top priority for operators, as its achievement leads to some significant reduction of CAPEX and OPEX, thus eventually increasing the operating revenues. Let us boil down Problem 1 to two constituent fundamental problems, namely Throughput Problem and Positioning Problem. First, we aim at maximizing the throughput fairness among test points while ensuring that localization performances are below a given threshold ζ_b . This threshold is a key parameter of LOKO and is thoroughly analysed in Section 5.4. The above problem can be cast into the following

Problem 2 (Throughput Problem).

$$\max_{\mathbf{A}, \mathbf{x}} \quad \min_{t \in \mathcal{T}} r \quad (14a)$$

$$\text{subject to} \quad \mathbf{b} \leq \zeta_b^2 \mathbf{1}_T, \quad (14b)$$

$$(10b), (10c), (10d), (10e). \quad (14c)$$

Note that $\zeta_b \leq u_t \forall t$, as LTE positioning performances represent a baseline for our solution.

Positioning problem. Conversely to the previous problem, one operator may optimize the positioning performances of the network while satisfying some throughput performance threshold ζ_r . Again, we take the LTE network performances as baseline by assuming $\zeta_r \geq r_t \forall t$. Therefore, we can define the positioning problem as follows

Problem 3 (Positioning Problem).

$$\min_{\mathbf{A}, \mathbf{x}} \quad \max_{t \in \mathcal{T}} \mathbf{b} \quad (15a)$$

$$\text{subject to} \quad \mathbf{r} \geq \zeta_r \mathbf{1}_T, \quad (15b)$$

$$(10b), (10c), (10d), (10e). \quad (15c)$$

The two problems are non-convex, thus hard to solve as they cannot be tackled with convex optimization techniques. We discuss their non-convexity property in the next section while developing suitable algorithmic solutions.

5 CONVEX RELAXATION AND SOLUTION

In order to provide a practical and affordable solution, we apply a convex relaxation to the above-described problems.

5.1 Convexity Analysis

We first prove that Problems 2 and 3 are non-convex. Since \mathbf{r} and \mathbf{b} swap roles in the problems being objective function and constraint, respectively, we can focus on one single problem, i.e., Problem 2, leaving the non-convexity property of Problem 3 as an analogous proof.

Algorithm 1: Double-nested BCA algorithm for Problem 2

```

Initialize  $\mathbf{A}, \mathbf{x}, \mathbf{y}$  to feasible values  $\hat{\mathbf{A}}^{(k=0)}, \bar{\mathbf{x}}^{(k=0)}$ 
repeat                                     Outer BCA loop (over  $k$ )
  repeat                                   Inner BCA loop (over  $i$ )
    Derive  $\bar{\mathbf{y}}^{(i)}$  by solving Problem 2 as  $\max_{\mathbf{y}} \min_{t \in \mathcal{T}} \mathbf{r}''$ ,
      where  $\mathbf{r}''$  is obtained from Eq. (19)
    Derive  $\bar{\mathbf{X}}^{(i)}$  by solving Problem 4
    Derive  $\bar{\mathbf{x}}^{(i)}$  by applying the Gaussian
      Randomization to  $\bar{\mathbf{X}}^{(i)}$ 
  until convergence of  $\mathbf{r}'$  as per Eq. (16)
  Set  $\hat{\mathbf{x}}^{(k)}$  to  $\bar{\mathbf{x}}^{(i)}$ 
  Derive  $\hat{\mathbf{A}}^{(k)}$  by solving Problem 2 given  $\hat{\mathbf{x}}^{(k)}$ 
until convergence of objective function Eq. (14a) in
Problem 2

```

Let us consider one element of \mathbf{r} , as per Eq. (11). The sum of ratios (of g'_{tj} that is function of x_j) shows the non-convexity in the optimization variable \mathbf{x} . Let us now consider the set of constraints. While constraints (10b), (10c) and (10d) are linear (and convex), constraints (14b) and (10e) are not. With regard to constraint (14b), b_t provides a linear term $a'_t u_t^2$, whereas the remaining one is a product of a linear function of a_{tj} and a non-convex function of x_j , thus making the constraint non-convex. Finally, constraint (10e) is non-convex as it enforces the binary nature of the decision variables. First off, we relax this constraint making a_{tj} and x_j continuous in the closed convex set $[0, 1]$. Then, we introduce a number of convex relation techniques and optimization methods tailored to address each non-convexity issue of the above-mentioned problems, and provide practical solutions.

5.2 Throughput Routine

We describe the devised routine used to solve Problem 2. It requires joint gNB associations and deployment decisions, while taking into account the throughput provided by the pre-existing LTE deployment. Note that one gNB is deployed at CS j if there exists a test point t that would get higher throughput w.r.t. the LTE network, given the optimal deployment decisions at the other sites, the maximized fairness among test points and the satisfaction of the positioning constraint. Besides, the total number of deployed gNBs must not exceed the allocated budget.

Let us consider the continuous relaxation of the problem, as already mentioned in Section 5.1. In the following, we develop a double-nested Block Coordinate Ascend (BCA) algorithm [27] as detailed in Algorithm 1, wherein association matrix \mathbf{A} is optimized while keeping decision variable \mathbf{x} fixed and vice-versa, in an iterative fashion¹² whose convergence proof is provided in the following.

Inner BCA loop. Let us start by selecting a feasible solution $\hat{\mathbf{A}}^{(k=0)}$ and solve the problem for \mathbf{x} . In this condition, the second addend in every element of \mathbf{r} is non-negative and independent of \mathbf{x} as per Eq. (11), so it can be neglected during this optimization stage. Therefore, we can re-write \mathbf{r} in Problem 2 as \mathbf{r}' , whose elements are defined as

$$r'_t = \sum_{j \in \mathcal{S}} \hat{a}_{tj}^{(k)} \log_2 \left(1 + \frac{g_{tj}}{g'_{tj} + N'} \right), \quad (16)$$

12. To avoid notation clutter, we annotate variables computed within the inner and outer loop with $\bar{\cdot}$ and $\hat{\cdot}$, respectively.

at the k th outer iteration, where we drop the bandwidth W_g , being it only a scale factor. We tackle the above problem by decomposing it into a sequence of convex problems. First, let us define a sum-of-functions-of-ratios problem as

$$\max_{\mathbf{z}} \quad \min_t \sum_j f_{tj} \left(\frac{\Lambda_{tj}(\mathbf{z})}{D_{tj}(\mathbf{z})} \right) \quad (17a)$$

$$\text{subject to} \quad \mathbf{z} \in \mathcal{Z}, \quad (17b)$$

where functions $f_{tj}(\cdot)$ are non-decreasing, $\Lambda_{tj}(\cdot)$ are non-negative, $D_{tj}(\cdot)$ are positive and \mathcal{Z} is any non-empty real set. By involving a sum of ratios, the above problem is a multiple-ratio Fractional Programming (FP) problem. Classical FP techniques such as Charnes-Cooper Transform or Dinkelbach's Transform cannot be readily implemented in the context of multiple-ratio FP as they would not guarantee the equality between the optimal transformed and original objective function values. Therefore, to decouple the numerator and the denominator of each addend of the fractional objective function, we can leverage on the *Quadratic Transform* [28], which leads to the following equivalent formulation

$$\max_{\mathbf{z}, \mathbf{y}} \quad \min_t \sum_j f_{tj} (2y_{tj} \sqrt{\Lambda_{tj}(\mathbf{z})} - y_{tj}^2 D_{tj}(\mathbf{z})) \quad (18a)$$

$$\text{subject to} \quad y_{tj} \in \mathbb{R}, \quad \forall t \in \mathcal{T}, j \in \mathcal{S}, \quad (18b)$$

$$\mathbf{z} \in \mathcal{Z}, \quad (18c)$$

where \mathbf{y} denotes the collection of auxiliary variables y_{tj} for each fractional term $\Lambda_{tj}(\mathbf{z})/D_{tj}(\mathbf{z})$. We consider a particular type of FP problem, namely concave-convex, by introducing some additional hypotheses: *i*) $f_{tj}(\cdot)$ and $\Lambda_{tj}(\cdot)$ are concave functions, *ii*) $D_{tj}(\cdot)$ are convex functions, *iii*) \mathcal{Z} is a nonempty convex set expressed by a finite number of inequality constraints, i.e. in standard form. In these conditions, we can solve the problem for the primary variable \mathbf{z} and the auxiliary variable \mathbf{y} iteratively, while dealing with a convex problem at every stage by holding one variable fixed at a time. Specifically, the objective function in Eq. (16) produces a concave-convex FP as $\log(\cdot)$ and $g_{tj}(\cdot)$ are concave functions and $g'_{tj}(\cdot)$ are convex functions in their domain. Thus, we proceed by developing an inner BCA loop (in contrast to the outer one addressing the optimization of variables \mathbf{x} and \mathbf{A}), as shown in Algorithm 1. To this end, we apply the above-mentioned Quadratic Transform to the objective function in Eq. (16) for the k th step and obtain the following

$$r_t'' = \sum_{j \in \mathcal{S}} \hat{a}_{tj}^{(k)} \log_2 \left(2y_{tj} \sqrt{g_{tj} + g'_{tj} + N'} - y_{tj}^2 (g'_{tj} + N') \right), \quad (19)$$

which enables solving the problem for the primary variable \mathbf{x} and the auxiliary variable \mathbf{y} iteratively while holding one variable fixed at a time. Indeed, the problem is now convex in \mathbf{y} for a given \mathbf{x} , thus solvable by means of convex programming, providing the optimal solution $\bar{\mathbf{y}}$. By substituting such value into Eq. (19), we need to solve the problem for \mathbf{x} , which is still non-convex in \mathbf{x} due to nature of \mathbf{b} in constraint (14b). Hereafter, we take advantage of the Semidefinite Relaxation (SDR) technique to relax such constraint [29]. Note that, by doing so, we derive a fully convex relaxation also of the problem in \mathbf{x} , thereby ensuring the convergence of the inner BCA loop [27]. As a preliminary step, we can re-write such a constraint as follows

$$(1 - \hat{a}_t'^{(k)}) \boldsymbol{\nu}_t^\top \mathbf{x} \leq (\zeta_b^2 - \hat{a}_t'^{(k)} u_t^2) \mathbf{x}^\top \mathbf{F}_t^\top \mathbf{x}, \quad (20)$$

given that the quadratic form $\mathbf{x}^\top \mathbf{F}_t^\top \mathbf{x}$ is non-zero. This condition holds if there exist at least two non-zero elements of \mathbf{x} whose corresponding sites do not share the same angular position with respect to CS t , which would set the squared sine of their difference to zero¹³. Moreover, there is no guarantee that the symmetrisation of \mathbf{F}_t is negative semidefinite (NSD), hence Eq. (20) is yet non-convex.

The SDR procedure stems from the observation that $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ if and only if $X \succeq 0$ and $\text{rank}(\mathbf{X}) = 1$. This observation allows linearizing the constraint as a function of \mathbf{X} , rather than \mathbf{x} . SDR drops the non-convex constraint $\text{rank}(\mathbf{X}) = 1$ in order to obtain a convex relaxation of the original problem. To this aim, we point out that $\mathbf{x}^\top \mathbf{F}_t^\top \mathbf{x} = \text{Tr}(\mathbf{F}_t^\top \mathbf{X})$ as per a well-known matrix identity, while $\boldsymbol{\nu}_t^\top \mathbf{x} = \text{Tr}(\mathbf{X} \odot \mathbf{V}_t)$ due to the original binary nature of \mathbf{x} ¹⁴, where \mathbf{V}_t is a diagonal matrix such that $\text{diag}(\mathbf{V}_t) = \boldsymbol{\nu}_t$. Therefore, we can write the following problem:

Problem 4 (Inner BCA stage for \mathbf{x}).

$$\max_{\mathbf{X}} \quad \min_{t \in \mathcal{T}} r'' \quad (21a)$$

$$\text{subject to} \quad \hat{\mathbf{A}}^{(k)} \leq \mathbf{1}_T \text{diag}(\mathbf{X})^\top, \quad (21b)$$

$$(1 - \hat{a}_t'^{(k)}) \text{Tr}(\mathbf{X} \odot \mathbf{V}_t) \leq (\zeta_b^2 - \hat{a}_t'^{(k)} u_t^2) \text{Tr}(\mathbf{F}_t^\top \mathbf{X}), \quad \forall t \in \mathcal{T}, \quad (21c)$$

$$\text{Tr}(\mathbf{X}) = G, \quad (21d)$$

$$\mathbf{X} \succeq 0, \quad (21e)$$

$$\text{diag}(\mathbf{X}) \in [0, 1]^{\mathcal{S}}, \quad (21f)$$

where we amend g'_{tj} in Eq. (19) as $g_{tj}'' = \sum_{n \in \mathcal{S} \setminus \{j\}} g_{tn} X_{nn}$ to comply with the new optimization variable \mathbf{X} , and re-write constraints (10b) and (10d) as (21b) and (21d), respectively.

The above problem can be solved by the interior-point method. Let us denote the solution of the i th step as $\{\tilde{\mathbf{x}}^{(i)}, \tilde{\mathbf{X}}^{(i)}\}$. Due to the SDR, the matrix $\tilde{\mathbf{X}}^{(i)}$ is not necessarily rank-1. If $\text{rank}(\tilde{\mathbf{X}}^{(i)}) = 1$, then we round $\tilde{\mathbf{x}}^{(i)}$ by setting its G largest elements to 1 and the remaining to 0, and check if the result $\bar{\mathbf{x}}^{(i)}$ is feasible for the non-relaxed problem, i.e. for the throughput problem stage for \mathbf{x} after the quadratic transform. If $\text{rank}(\tilde{\mathbf{X}}^{(i)}) > 1$ or the obtained $\bar{\mathbf{x}}^{(i)}$ is not feasible for the non-relaxed problem, we leverage on the Gaussian Randomization procedure to extract a feasible rank-1 component from $\tilde{\mathbf{X}}^{(i)}$. Specifically, we generate a set of $L > 0$ random vectors $\boldsymbol{\xi}_l$, with $l = 1, \dots, L$, drawn from a Gaussian distribution with mean $\tilde{\mathbf{x}}^{(i)}$ and covariance matrix $\tilde{\mathbf{X}}^{(i)}$, i.e. $\boldsymbol{\xi}_l \sim \mathcal{N}(\tilde{\mathbf{x}}^{(i)}, \tilde{\mathbf{X}}^{(i)})$, and project them onto the feasible region of the non-relaxed problem by setting the G largest elements to 1 and the rest to 0 for each random sample. Then, to enforce the PEB constraint in Eq. (14b), we remove each random sample that does not satisfy such constraint, yielding a sample size $L' \leq L$. Lastly, we choose among the obtained L' feasible solutions the one $\bar{\mathbf{x}}$ that maximizes the objective function (19) to be the rank-1 solution of Problem 4.

We keep executing the inner BCA loop and solve for \mathbf{y} and \mathbf{x} , obtaining $\bar{\mathbf{y}}^{(i)}$ and $\bar{\mathbf{x}}^{(i)}$ at the i th inner iteration. We denote the resulting solution for \mathbf{x} after convergence is reached as $\hat{\mathbf{x}}^{(k)}$, being k the current outer BCA iteration.

13. The optimal solution naturally avoids such pathological conditions as the resulting PEB would explode, thus not satisfying the corresponding constraint.

14. Due to the original binary constraint (10e), $x_j = x_j^2$. As $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$, we have that $X_{jj} = x_j$.

Algorithm 2: BCD algorithm for Problem 3

```

Initialize  $\mathbf{A}$  to feasible values  $\hat{\mathbf{A}}^{(k=0)}$ ,  $0 < \eta_l < \eta_u$  and  $\epsilon > 0$ 
repeat
  Set  $\eta$  to  $(\eta_l + \eta_u)/2$ 
  repeat
    if Problem 5 is feasible then
      | Set  $\eta_u$  to  $\eta$ , Set  $\bar{\mathbf{x}}^{(i)}$  to its solution
    else
      | Set  $\eta_l$  to  $\eta$ 
    end
  until  $\eta_u - \eta_l \leq \epsilon$ 
  Set  $\hat{\mathbf{x}}^{(k)}$  to  $\bar{\mathbf{x}}^{(i)}$ 
  Derive  $\hat{\mathbf{A}}^{(k)}$  by solving Problem 3 given  $\hat{\mathbf{x}}^{(k)}$ 
until convergence of objective function (15a) in Problem 3

```

Outer BCA loop. Conversely, let us hold \mathbf{x} fixed to the optimal value $\hat{\mathbf{x}}^{(k)}$ and solve Problem 2 for \mathbf{A} (the problem is now linear). We keep solving for \mathbf{x} and \mathbf{A} , iterating the two outer optimization stages until convergence, and indicate the current optimal solutions as $\hat{\mathbf{x}}^{(k)}$ and $\hat{\mathbf{A}}^{(k)}$, respectively.

Proposition 1. The double-nested BCA algorithm for Problem 2 converges to a stationary point.

Proof: According to [27], the convergence of BCA is guaranteed under the condition that the objective function and the feasible set are convex in each block of variables, although non-convex in general. Indeed, by means of the Quadratic Transform, the Semidefinite Relaxation (SDR) and the continuous relaxation of the binary constraints, we derive a convex relaxation of Problem 2. As described above, such formulation is convex in \mathbf{A} , \mathbf{x} and \mathbf{y} respectively, thus guaranteeing the convergence of the inner BCA loop at each outer loop iteration and then the eventual convergence of the outer BCA loop.

5.3 Positioning Routine

We study Problem 3 by partially retracing the steps performed to solve Problem 2 and we propose an analogous routine. In particular, we make use of the Bisection Method [30] within an analogous Block Coordinate Descent (BCD) algorithm while restating constraint (14b) in an equivalent form. Indeed, straightforward algebraic manipulations show that the latter is equivalent to the following set of constraints

$$\hat{a}_{tj}^{(k)}(g'_{tj} + N')(2^{\zeta_r/W_g} - 1) \leq g_{tj} \quad \forall t \in \mathcal{T}, \forall j \in \mathcal{S} \quad (22)$$

$$\hat{a}_t^{(k)} \left(\frac{\zeta_r}{W_e} \right) \leq c_t \quad \forall t \in \mathcal{T}, \quad (23)$$

out of which Eq. (22) is notably still non-convex since it contains a product of optimization variables, being $g'_{tj} = \sum_{n \in \mathcal{S} \setminus \{j\}} g_{tn} x_n$ as defined in Section 4.1. Both loops are detailed in Algorithm 2.

Inner Bisection loop. Let us describe the inner Bisection loop¹⁵ by picking a feasible initial solution $\hat{\mathbf{A}}^{(k=0)}$ for the outer

15. Note that performing the Quadratic Transform leads to minimizing the maximum of a concave objective function, which is yet a non-convex problem.

Algorithm 3: Adaptive KPI Processing

```

Initialize  $\zeta_b \gg \max_t u_t$ ,  $\delta$  to feasible value based on  $\mathcal{S}$  and  $\epsilon > 0$ 
Obtain  $r^{(\tau=0)}$ ,  $\hat{\mathbf{x}}^{(\tau=0)}$ ,  $\hat{\mathbf{A}}^{(\tau=0)}$  via Algorithm 1
Derive  $b^{(\tau=0)}$  via Eq. (13)
repeat
  Set  $\omega_b^{(\tau)}$  to  $\frac{b^{(\tau-1)} - 1}{b^{(\tau-1)}}$ ,  $\zeta_b^{(\tau)} = \omega_b^{(\tau)} b^{(0)}$ 
  Obtain  $r^{(\tau)}$  via Algorithm 1
  Set  $\omega_r^{(\tau)}$  to  $\frac{r^{(\tau-1)} - \mu}{r^{(\tau-1)}}$ ,  $\zeta_r^{(\tau)}$  to  $\omega_r^{(\tau)} r^{(0)}$ 
  Obtain  $\hat{b}^{(\tau)}$ ,  $\hat{\mathbf{x}}^{(\tau)}$ ,  $\hat{\mathbf{A}}^{(\tau)}$  via Algorithm 2
  Derive  $\mathbf{x}^{*(\tau)}$  by rounding  $\hat{\mathbf{x}}^{(\tau)}$ 
  for  $j \in \mathcal{S}$ ,  $t \in \mathcal{T}$  do
    if  $x_j^* = 1$  &  $\hat{a}_{tj} \leq \delta \max_t \hat{a}_{tj}$  then
      | Set  $a_{tj}^*$  to 0
    else if  $x_j^* = 0$  then
      | Set  $a_{tj}^*$  to 0
    end
  end
  for  $t \in \mathcal{T}$  do
    Set  $a_{tj}^*$  to 1 for  $j = \arg \max_{n | \hat{a}_{tn} \neq 0} \frac{g_{tn}}{g'_{tn}(\mathbf{x}^*) + N'}$ 
  end
  Derive  $r^{*(\tau)}$  via Eq. (11) and  $b^{*(\tau)}$  via Eq. (13)
until Overall LOKO convergence  $\left| \frac{r^{(0)} - r^{*(\tau)}}{b^{(0)} - b^{*(\tau)}} - \mu \right| \leq \epsilon$ 

```

loop. While holding $\hat{a}_t^{(k)}$ fixed, the second term of \mathbf{b} is non-negative and independent of \mathbf{x} , and so it can be neglected. Indeed, we can re-write the t th element of \mathbf{b}' as

$$b'_t = (1 - \hat{a}_t^{(k)}) \frac{\mathbf{v}_t^\top \mathbf{x}}{\mathbf{x}^\top \mathbf{F}_t^\top \mathbf{x}}, \quad (24)$$

being the denominator $\mathbf{x}^\top \mathbf{F}_t^\top \mathbf{x}$ non-negative. However, as discussed in Section 4.2, the quadratic form $\mathbf{x}^\top \mathbf{F}_t^\top \mathbf{x}$ is non-convex, thus calling for the usage of SDR. The resulting function is now quasi-convex and can be tackled by means of the Bisection method. Furthermore, by keeping fixed either \mathbf{x} or \mathbf{A} , constraints (22) and (23) are convex, leading us to state the following feasibility problem, which enables solving the inner Bisection stage for \mathbf{x} , namely

Problem 5 (Inner Bisection stage for \mathbf{x}).

$$\begin{aligned} \text{Find} \quad & \mathbf{X} & (25a) \\ \text{subject to} \quad & (1 - \hat{a}_t^{(k)}) \text{Tr}(\mathbf{X} \odot \mathbf{V}_t) \leq \eta \text{Tr}(\mathbf{F}_t^\top \mathbf{X}), \quad \forall t \in \mathcal{T}, & (25b) \end{aligned}$$

$$(21b), (21e), (21d), (21f), (22)'$$

where \mathbf{X} and \mathbf{V}_t attain to the same notation used in Section 5.2 for the Throughput Problem, $\eta \in [\eta_l, \eta_u] \subset \mathbb{R}_+$ is the bisection threshold and constraint (22)' is the constraint in Eq. (22) compliant with the SDR formulation, by substituting g'_{tj} with $g''_{tj} = \sum_{n \in \mathcal{S} \setminus \{j\}} g_{tn} X_{nn}$. Note that the Bisection Method requires exactly $\log_2((\eta_u - \eta_l)/\epsilon)$ iterations, where $\epsilon > 0$ is a given tolerance. We apply the same Gaussian Randomization procedure as aforementioned in Section 4.2 that returns a rank-1 solution $\bar{\mathbf{x}}^{(i)}$. We keep checking the feasibility of the inner Bisection stage while decreasing η at each iteration until $\eta_u - \eta_l \leq \epsilon$. We denote the feasible solutions provided at the i th iteration as $\bar{\mathbf{y}}^{(i)}$ and $\bar{\mathbf{x}}^{(i)}$, respectively.

TABLE 2: Simulation parameters.

	σ_0	λ [m]	P [W]	f [GHz]	W [MHz]
Legacy LTE	10^{-3}	10	30	1.8	20
5G-NR	10^{-4}	1	20	3.5	100

Outer BCD loop. Besides, we indicate the inner loop solution for \mathbf{x} upon convergence as $\hat{\mathbf{x}}^{(k)}$, where k is the outer iteration. Then, we plug it into the problem for \mathbf{A} , now linear with constraints (22) and (23), and obtain the solution $\hat{\mathbf{A}}^{(k)}$.

Proposition 2. The BCD algorithm for Problem 3 converges to a stationary point.

Proof: The proof of this proposition goes analogously as the one provided for Proposition 1.

5.4 Adaptive KPI Processing and Overall Convergence

The Adaptive KPI Processing module returns thresholds ζ_r and ζ_b for each LOKO iteration τ , which consists of executing both Throughput and Positioning Routines, and the quantization of relaxed variables, as clarified in Algorithm 3. Binary \mathbf{x}^* is obtained by rounding $\hat{\mathbf{x}}$, while binary \mathbf{a}^* is derived via the following steps in order to satisfy the constraint in Eq. (10c): *i*) set \hat{a}_{tn} such that $\hat{a}_{tn} \leq \delta \max_t \hat{a}_{tn}$ with $n \in \{j|x_j = 1\}$ or every \hat{a}_{tn} with $n \in \{j|x_j = 0\}$ to zero, where δ is finely adjusted based on S , *ii*) $\forall t$, set the element $\hat{a}_{tj} \neq 0$ providing the highest ratio $g_{tj}/(g'_{tj} + N')$ (given \mathbf{x}^*) to one.

As the two routines pursue conflicting objectives in some configurations, it is unlikely that the output of one routine plugged in the other as threshold would lead to a feasible solution, thus calling for a scaling factor finely tuned to ensure the feasibility of each problem in the way to convergence. First, we run the Throughput Routine with $\zeta_b \gg \max_t u_t$, obtaining its solution $r^{(\tau=0)}$ along with $\mathbf{x}^{(\tau=0)}$ and $\mathbf{A}^{(\tau=0)}$ that in turn provide $b^{(\tau=0)}$ via Eq. (13). Then, we set $\zeta_r^{(\tau)} = \omega_r^{(\tau)} r^{(0)}$, with $\omega_r^{(\tau)} = \frac{r^{(\tau-1)} - \mu}{r^{(\tau-1)}}$, and execute the Positioning Routine to get $b^{(\tau)}$, while the quantization procedure provides $r^{*(\tau)}$ and $b^{*(\tau)}$. If $\frac{r^{(0)} - r^{*(\tau)}}{b^{(0)} - b^{*(\tau)}} \simeq \mu$, convergence is reached. Conversely, we set $\zeta_b^{(\tau)} = \omega_b^{(\tau)} b^{(0)}$, with $\omega_b^{(\tau)} = \frac{b^{(\tau-1)} - 1}{b^{(\tau-1)}}$ and keep iterating until convergence.

6 NUMERICAL RESULTS

We perform an extensive simulation campaign to test LOKO performances: we first validate our solution via Monte Carlo simulations on large instances by means of synthetic traces derived from realistic deployment settings. We then benchmark LOKO against the closest available state-of-the-art literature on a real dense-urban network topology (Bologna, Italy) provided by a major European network operator, wherein legacy (LTE) base stations are already in place and 5G-NR candidate sites are properly chosen on the map. In the following, we use terms Rate and Throughput, interchangeably. Simulation parameters are listed in Table 2.

6.1 Synthetic Topologies

We consider three different scenarios, namely Highway (H), Suburban (SU) and Dense-Urban (DU), where we place legacy base stations (eNBs) and 5G-NR (gNB) candidate sites (CSs). Examples are depicted in Fig. 2. For each simulation execution, we consider no deployment budget limitation, i.e. $G = S$, and randomly place eNBs and CSs while keeping fixed the base station

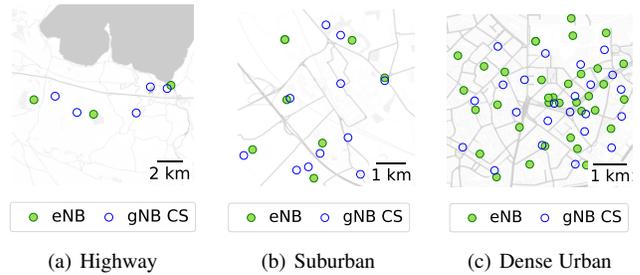


Fig. 2: Considered scenarios based on real topologies.

TABLE 3: Parameters by scenario.

	Legacy LTE			5G-NR		
	H	SU	DU	H	SU	DU
α	2.5	3	3.5	2.5	3	3.5
σ_s [dB]	3	5	6	5	7	9

density values obtained through realistic topologies of a major European Operator. Besides, we set propagation factor α and log-normal shadowing standard deviation σ_s for each scenario as per Table 3.

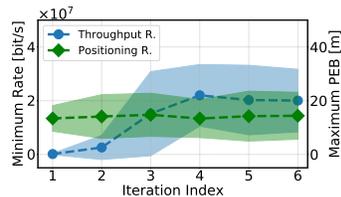


Fig. 3: Routines convergence average and standard deviation (shaded) in DU scenario.

Scenario	Throughput Routine cycles	Positioning Routine cycles
H	2	2
SU	3	2
DU	5	4

TABLE 4: Routines convergence times.

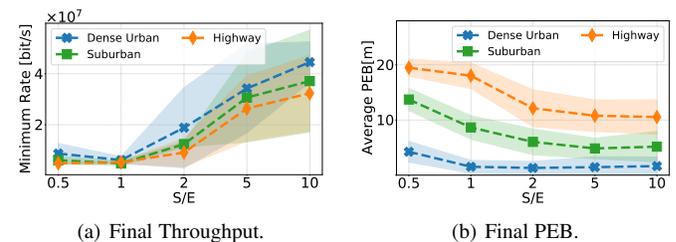
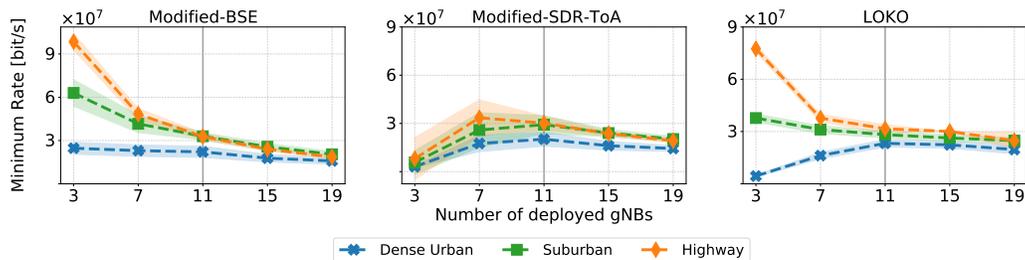
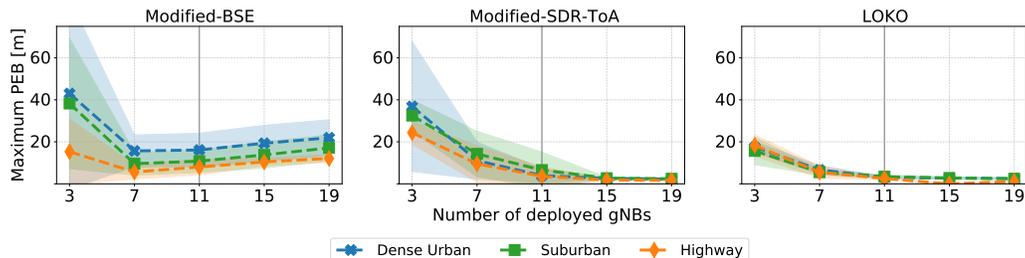


Fig. 4: LOKO performances and corresponding standard deviation (shaded) for different CS densities.

Convergence analysis and time complexity. Given the relaxation procedure described in Section 5, we first validate the convergence of each individual routine. Fig. 3 illustrates the evolution of each routine over time, showing that they converge after a few iterations. In Table 4, we list the convergence time of each routine for the various considered scenarios. As expected, while the number of eNBs and available CSs grows (see e.g., Dense-Urban scenario) the proposed algorithm takes slightly longer to converge. In particular, the complexity of LOKO is driven by the interior point problem in each routine, which can be solved by the interior point method with a worst-case complexity of $\mathcal{O}(S^{5.5})$ according to [31], [32]. Therefore, LOKO has greatly lower complexity than exhaustive search, which is in the order of $\mathcal{O}(S \sum_{i,j}^{S+1} \binom{T}{i} \binom{S+1}{j})$. While rapid execution might be beneficial for the computational affordability of LOKO, the main focus of our derivation is to



(a) LOKO performance in terms of minimum rate against Modified-BSE and Modified-SDR-ToA.



(b) LOKO performance in terms of maximum PEB against Modified-BSE and Modified-SDR-ToA.

Fig. 5: LOKO performance against Modified-BSE and Modified-SDR-ToA for different numbers of deployed gNBs. The vertical line indicates the number of deployed gNBs to preserve the deployment density of the reference European major network operator in each simulation scenario.

provide a solution proved to converge regardless of the number of iterations required.

Large-instances evaluation. We analyze now the overall performances of LOKO increasing the ratio between available CSs and eNBs (S/E) with no deployment budget limitation, namely with $G = S$. When not explicitly mentioned, we set $\text{TPR } \mu = 10$. Moreover, we show the average PEB to reflect the maximum positioning accuracy experienced by users in the area. In Fig. 4, we show different performances for each selected scenario in terms of minimum throughput and average PEB. When the number of CSs increases, more degrees of freedom are introduced thus a higher granularity can be reached. However, this might require a higher CAPEX for operators that may need to install a larger number of gNBs.

Effect of the number of deployed gNBs. Hereafter we investigate the behavior of LOKO with respect to the number of deployed gNBs G compared against the above-mentioned state-of-the-art techniques. As supported in Section 7, this is the first work of its kind addressing the cellular network planning problem while optimizing user throughput, fairness and positioning KPIs. Nonetheless, to provide a significant comparison against the available literature, we hand-pick two state-of-the-art solutions, which seek the independent optimization of network throughput and localization accuracy, respectively.

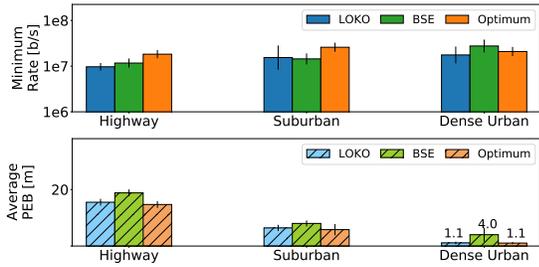
The former is dubbed as Base Station Elimination (BSE) and aims at minimizing the number of base stations covering a target area while satisfying the throughput requirements [33]. By default, BSE does not support multiple RATs nor a fixed number G of gNBs to be deployed. Therefore, we perform some modifications of the original algorithm and design a *Modified-BSE* to account for two non-interfering RATs as well as to control the number of deployed BSs. In particular, we begin with a full gNBs deployment and iteratively discard the candidate site whose corresponding gNB produces the highest total SINR (i.e., sum SINR over all test points) when eliminated in comparison with removing any other

candidate site. The algorithm stops when the required number of gNBs is deployed, converging in exactly $S - G$ iterations.

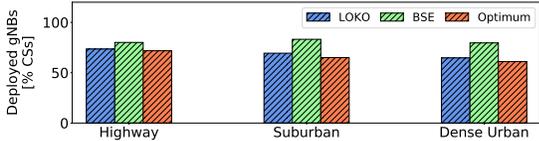
The latter benchmark in [34] addresses the problem of selecting the set of anchor nodes in a dense sensor network deployment delivering the best localization performance at a given user location. Such algorithm achieves nearly optimal sensor selection by minimizing the trace of the CRLB, whose direct relationship with the PEB is described in Section 2.2, under the condition of fixed number of selected sensors. As the original formulation is not bespoke for the purpose of cellular network deployment, we introduce a slight modification of its objective, namely we optimize the PEB fairness among all UEs locations by minimizing the maximum of the above-mentioned CRLB trace over all test points. It is worth pointing out that this solution complies with the widely used simplifying assumption that ToA measurement errors are independent of the actual distance, which translates into a fixed-variance Gaussian noise. Therefore, it optimizes for a PEB formulation that accounts only for the angular positions of the CSs with respect to the test points, which is not guaranteed to provide good performance when evaluating the resulting PEB as per Eq. (13). We refer to this amended version of the algorithm as *Modified-SDR-ToA* approach.¹⁶ Furthermore, the exhaustive search method is introduced in the performance benchmark as a greedy approach if computationally affordable for the specific problem instance.

Fig. 5 illustrates the throughput and localization performances of LOKO compared against the above-mentioned benchmarks for different numbers of deployed gNBs. We consider a range of gNBs such that the middle point $G = 11$, indicated by a vertical line, returns the same BS density as the one provided by our reference major European operator. As for the minimum rate depicted in Fig. 5(a), Modified-BSE has a monotonic decreasing behavior

¹⁶ Modified-SDR-ToA has a complexity of $\mathcal{O}(S^{4.5})$ while BSE has a complexity of $S^2\mathcal{O}((S-1)T)$.



(a) LOKO performance against BSE for the same number of deployed gNBs as obtained at BSE convergence



(b) Number of deployed gNBs by LOKO against BSE when delivering the same average PEB performance as BSE upon its convergence.

Fig. 6: LOKO CAPEX efficiency against the BSE algorithm (c.f. [33]) and the optimum derived via exhaustive search.

while Modified-SDR-ToA has an approximately concave trend. Indeed, the latter technique attempts at optimally selecting candidate sites in the service area so as to improve upon the minimum PEB provided by the LTE network. However, the 5G network proves incapable to deliver better localization performance than LTE in case only a handful of gNBs are deployed, thereby forcing several test points to be served by the LTE network under the assumption of hybrid BS-UE association, as illustrated in Section 4.1. Increasing the number of gNBs, the 5G deployment starts outmatching the LTE PEB and moves more and more test points to the 5G network, resulting in a bump in throughput performance in the middle gNBs range. Notably, LOKO has a cross-breed behavior, which traces Modified-BSE in the highway and suburban scenarios, and Modified-SDR-ToA in the urban scenario, in the low number of deployed gNBs regime. In particular, LOKO is able to deliver higher throughput with a few available gNBs whenever the PEB constraint is comparatively easier to satisfy, e.g., in the H and SU scenarios, thanks to the inherently higher spatial sparsity of BSs with respect to the test points in bigger deployment areas.

On the other hand, the analysis in Fig. 5(b) of the localization performance shows the great advantage of LOKO against its competitor solutions. As expected, Modified-BSE provides the worst results but even Modified-SDR-ToA is outmatched by as much as 60% (reduction in PEB w.r.t. the value obtained by Modified-SDR-ToA in low gNBs regime). The reason behind it is that Modified-SDR-ToA does not consider the measurement error dependency on the actual distance from the anchor BS, therefore calculating sub-optimal UE-BS associations. We would like to highlight that LOKO delivers similar performance for all scenarios in terms of maximum PEB, which is likely provided by the LTE network. Nonetheless, these results prove that LOKO is able to deploy gNBs as well as establish UE-BS associations without letting the majority of UEs connect to the less-performing LTE network by providing an effective 5G deployment even with a few available gNBs to be deployed.

CAPEX utilization efficiency. Inspired by the original BSE [33], which does not assume any constraint on the number of gNBs, we test the efficiency of LOKO in terms of CAPEX utilization.

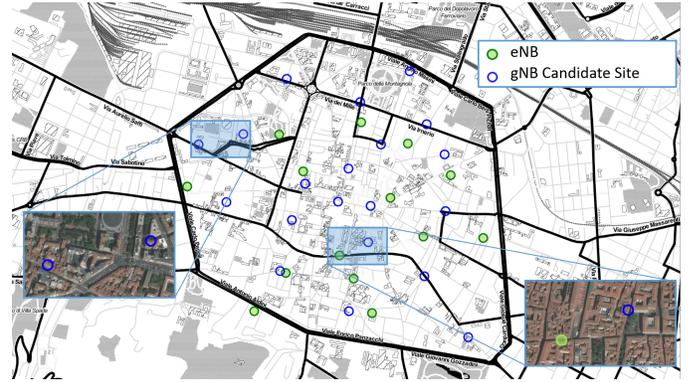


Fig. 7: Legacy (LTE) deployment and selected 5G-NR CSs.

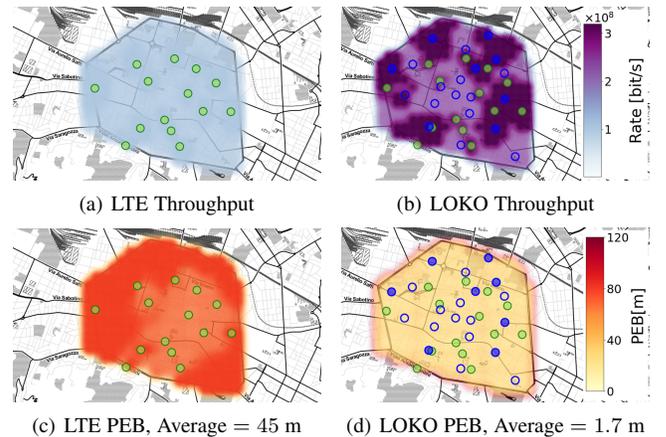


Fig. 8: Final results on Bologna Ringway.

Specifically, we consider the same scenarios (H, SU, DU) and run BSE until convergence. Then, we run LOKO and set G equal to the number of deployed gNBs upon BSE convergence. Note that BSE convergence happens at different number of iterations in each scenario. In Fig. 6(a), we depict the corresponding performance in terms of minimum rate and average PEB. Interestingly, LOKO significantly outperforms BSE (about -75%) in terms of PEB while showing similar throughput results (about -10% on average) when the same number of gNBs is utilized, achieving near-optimal results at an affordable time complexity. In addition, we report in Fig. 6(b) the fraction of deployed gNBs with respect to the overall CSs such that LOKO and BSE approximately deliver the same localization performance, proving that LOKO has a lighter footprint on the available resources in almost all tested scenarios. These results unveil that optimally planning the deployment at the available sites is crucial to pursue joint throughput and localization performance maximization. Moreover, they greatly motivate the adoption of LOKO for dense-urban scenarios where geographical constraints require a complex but practical solution to reach the expected performances.

6.2 Realistic Deployments

Dense-urban scenarios fully exploit the ability of our solution to correctly balance off between throughput and positioning performances. Thus, we use a live-network dataset from a major European operator considering eNBs deployed within the Bologna Ringway [35]. We add a new tier that includes 5G-NR CSs, prop-

TABLE 5: Final results on Bologna Ringway against TPR.

TPR	M-BSE	M-SDR-ToA	0	5	10	20
Min. Throughput [Mbit/s]	53.6	50.4	50.1	45.5	33.5	5.6
Maximum PEB [m]	18.6	16.7	14.1	12.9	12.4	11.9
Average PEB [m]	13.1	11.5	7.1	5.8	1.7	1.2

erly selected based on the operator guidelines from the Bologna map, as shown in Fig. 7.

We apply our solution to choose the minimum number of locations among 20 available CSs where to install gNBs. In Figs. 8(a) and 8(c), we show the performances of the legacy LTE deployment in terms of Throughput and PEB. LOKO deploys 8 gNBs and improves such performances by about 100% and 90% respectively, as shown in Figs. 8(b) and 8(d), partly due to 5G-NR boosted efficiency.

Finally, our solution enables operators to decide whether throughput performance has a higher priority w.r.t. PEB in the optimization process by means of TPR μ . Different values of TPRs and obtained results are summarized in Table 5. Note that when $\mu = 0$, LOKO optimizes only the minimum throughput whereas when $\mu = 20$ it mostly optimizes the PEB. We compare LOKO against Modified-BSE and Modified-SDR-ToA. Results show that LOKO already substantially outperforms Modified-BSE and Modified-SDR-ToA in terms of PEB even for $\mu = 0$, and further improves PEB as μ grows. As expected, throughput performance decreases as μ increases. For $\mu = 0$, LOKO provides a throughput performance comparable to BSE and M-SDR-ToA. As both Throughput and Positioning Routines need to be executed at least once before reaching convergence, even for $\mu = 0$, LOKO goes through the optimization of the PEB-constrained throughput and the throughput-constrained PEB, respectively. This has an impact on the final output, which results in a reduction of the minimum throughput of 0.6% w.r.t. M-SDR-ToA, which is well within the effect of numerical artifacts introduced by SDRs and related Gaussian randomization procedures. The deployment provided by M-SDR-ToA, by selecting CSs so as they are as much uniformly spread as possible around each test point, generates *as byproduct* good throughput performance in the considered dense urban environment, which is likely interference-limited. This is in full accordance with the results obtained on synthetic network topologies, as shown in Section 6.1.

7 RELATED WORK

Radio Frequency (RF) Planning. The problem of selecting where to install base stations is fundamental since the first cellular network deployments [36], [37]. The typical KPI for such works is coverage probability [38]–[40], which is shown to be almost invariant with respect to base station location geometry in case of massive deployments [41]. However, the advent of 5G and its advanced radio technology has led to a paradigm change [42]–[45] and brought up a series of new challenges, such as the usage of millimeter wave radio transceivers [46] and ultra-dense non-uniform deployment [47], [48] in which case a number of heuristic algorithms have been developed to optimize both coverage and throughput [49], [50]. To the best of our knowledge, this is the first work that considers localization accuracy in the network planning problem.

Localization Error bound. Providing a definition of localization error is crucial when dealing with positioning systems. The most used indicator is the CRLB, which has been derived neglecting measurement biases due to NLoS propagation conditions in early

works like [51], [52]. Conversely, more recent works address the measurement biases as additional estimation parameters [22], [53], which in turn improve the localization accuracy only if some prior information about them, e.g., their probability density function, is available [54], [55]. The adoption of mmWave radio technology has opened up new possibilities owing to their high channel angular sparsity, which allows estimating position-related parameters for each received NLoS path. For instance, [56] solves the UE localization problem for a multiple-input single-output (MISO) system, addressing the cases where massive arrays are deployed only at the base station side due to initial technical limitations integrating them into compact UEs. The bounds for 3D localization via mmWave are derived in [57], showing their theoretical capability of localizing a UE with sub-meter position error and sub-degree orientation error, though with differences in availability accuracy between uplink and downlink. Besides, mmWave MIMO can provide a progressive map of the environment by solving the so-called simultaneous localization and mapping (SLAM) problem, e.g. leveraging on diffuse multipath [58]. In our work, we have leveraged on a CRLB formulation that takes into account the dependency of the estimation error on the actual distance between the UE and the BS, thereby being general and able to capture the performance of all ToA-based techniques in the network planning phase.

Sensor Selection Problem. User positioning has been widely investigated in the literature and is still an active research track being of practical importance for many applications in wireless communications, ranging from navigation and radar systems to cellular and wireless sensor networks (WSNs) [59]–[61]. Specifically, it often translates to a sensor selection problem with different objectives, such as energy efficiency, connectivity robustness, positioning accuracy and more. Various works tackle these problems with heuristic approaches under the hypothesis of an uncorrelated or weakly correlated measurement model [62], [63], which can be efficiently handled by means of convex optimization. However, this assumption does not hold in real scenarios where sensor measurements are likely correlated [64]. Therefore, the resulting estimation accuracy is a non-linear function of the sensor selection variables [65], thus leading to a non-convex optimization formulation. Although greedy algorithms have been proposed [66], convex relaxation methods involving SDR provide greater performances, approaching exhaustive search at least on limited-size instances [34], [62], [67]. Besides the apparent difference in technological domain between WSNs and cellular networks, sensor selection problems do not generally aim at optimizing network throughput as the achievable communication rate is not a compelling parameter for sensor nodes. As a result, these works end up with problem formulations that are fundamentally different from ours.

To the best of our knowledge, there is not any prior literature addressing the joint optimization of network throughput and positioning performances in the network deployment process.

8 CONCLUSIONS

In this paper we proposed, analysed and evaluated LOKO, a novel method for 5G roll-out planning that, in contrast to state-of-the-art solutions, considers *mobile users localization accuracy requirements*. LOKO is designed to address the needs of service providers when considering *where to place 5G base stations*

to maximize throughput performances, taking into account *mobile users localization targets*. In order to accomplish this, we defined a new configuration parameter, *Throughput-Positioning Ratio (TPR)*, which enables operators to configure customized policies towards achieving a given positioning accuracy. LOKO iteratively relies on Block Coordinate, Semidefinite Relaxation (SDR) and Bisection techniques to find the right location for 5G base stations to meet predefined localization requirements within a given *available candidate sites set* and/or geographical area. Our results, considering *real base station deployments*, demonstrated the effectiveness of the solution along with the trade-offs between throughput and localization to be expected. Overall, in the large-scale scenarios considered, LOKO achieved up to one order of magnitude PEB improvements (tenths of meters to meters) depending on the *TPR* configuration.

9 ACKNOWLEDGEMENT

This work was supported by EU H2020 RISE-6G (grant agreement 101017011).

REFERENCES

- [1] P. Ahokangas *et al.*, “Business Models for Local 5G Micro Operators,” *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 730–740, 2019.
- [2] E. J. Oughton and Z. Frias, “The cost, coverage and rollout implications of 5G infrastructure in Britain,” *Telecommunications Policy*, vol. 42, no. 8, pp. 636–652, 2018.
- [3] 3GPP, “Study on NR positioning support,” Technical Report (TR) 38.855, 2019, version 16.0.0.
- [4] —, “Feasibility study on new services and markets technology enablers for critical communications; Stage 1,” Technical Report (TR) 22.862, 2014, version 14.1.0.
- [5] A. Albanese *et al.*, “SARDO: An Automated Search-and-Rescue Drone-based Solution for Victims Localization,” *IEEE Transactions on Mobile Computing*, 2021.
- [6] S. E. Elayoubi *et al.*, “5G RAN Slicing for Verticals: Enablers and Challenges,” *IEEE Comm. Magazine*, vol. 57, no. 1, pp. 28–34, 2019.
- [7] A. Albanese *et al.*, “First Responders Got Wings: UAVs to the Rescue of Localization Operations in Beyond 5G Systems,” *IEEE Communications Magazine*, vol. 59, no. 11, pp. 28–34, 2021.
- [8] J. Talvitie *et al.*, “Positioning and Location-Aware Communications for Modern Railways with 5G New Radio,” *IEEE Communications Magazine*, vol. 57, no. 9, pp. 24–30, 2019.
- [9] “R1-1901981 Further discussion of NR RAT-dependent UL Positioning,” 3GPP TSG-RAN WG1 Meeting #96, 2019.
- [10] Y. Liu *et al.*, “Prospective Positioning Architecture and Technologies in 5G Networks,” *IEEE Network*, vol. 31, no. 6, pp. 115–121, 2017.
- [11] P. Mursia *et al.*, “Risma: Reconfigurable intelligent surfaces enabling beamforming for IoT massive access,” *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 4, pp. 1072–1085, 2021.
- [12] H. Wymeersch *et al.*, “Radio Localization and Mapping With Reconfigurable Intelligent Surfaces: Challenges, Opportunities, and Research Directions,” *IEEE Vehicular Technology Magazine*, vol. 15, no. 4, pp. 52–61, 2020.
- [13] J. Schloemann *et al.*, “A Tractable Metric for Evaluating Base Station Geometries in Cellular Network Localization,” *IEEE Wireless Communications Letters*, vol. 5, no. 2, pp. 140–143, 2016.
- [14] M. Shafi *et al.*, “5G: A Tutorial Overview of Standards, Trials, Challenges, Deployment, and Practice,” *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1201–1221, 2017.
- [15] 3GPP, “Release 16 Description; Summary of Rel-16 Work Items,” Technical Report (TR) 21.916, 2020, version 0.4.0.
- [16] L. Wan *et al.*, “Enabling Efficient 5G NR and 4G LTE Coexistence,” *IEEE Wireless Communications*, vol. 26, no. 1, pp. 6–8, 2019.
- [17] European 5G Observatory, “USA, Japan, South Korea and China are the leading countries in terms of 5G readiness. A number of Gulf countries claim to have launched 5G mid-2018.” <https://5gobservatory.eu/5g-trial/major-international-5g-trials-and-pilots/>, 2020.
- [18] A. Dhekne *et al.*, “TrackIO: Tracking First Responders Inside-Out,” in *16th USENIX Symposium on Networked Systems Design and Implementation (NSDI 19)*. USENIX Association, Feb. 2019, pp. 751–764.
- [19] R. Kaune, “Accuracy studies for TDOA and TOA localization,” in *15th International Conference on Information Fusion*, 2012, pp. 408–415.
- [20] H. L. Van Trees and K. L. Bell, *On Geolocation Accuracy with Prior Information in Nonlineofight Environment*, 2007, pp. 558–561.
- [21] S. Fischer, “Observed time difference of arrival (OTDOA) positioning in 3GPP LTE,” *Qualcomm White Paper*, vol. 1, no. 1, pp. 1–62, 2014.
- [22] D. B. Jourdan *et al.*, “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.
- [23] D. B. Jourdan and N. Roy, “Optimal sensor placement for agent localization,” in *2006 IEEE/ION Position, Location, And Navigation Symposium*, Apr. 2006, pp. 128–139.
- [24] K. Yu *et al.*, “A Novel NLOS Mitigation Algorithm for UWB Localization in Harsh Indoor Environments,” *IEEE Transactions on Vehicular Technology*, vol. 68, no. 1, pp. 686–699, 2019.
- [25] S. M. Kay, *Fundamentals of Statistical Signal Processing : Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.
- [26] N. Levanon, “Lowest GDOP in 2-D scenarios,” *Radar, Sonar and Navigation, IEE Proceedings -*, vol. 147, pp. 149 – 155, 07 2000.
- [27] L. Grippo and M. Sciandrone, “On the Convergence of the Block Nonlinear Gauss-Seidel Method under Convex Constraints,” *Oper. Res. Lett.*, vol. 26, no. 3, p. 127136, Apr. 2000.
- [28] K. Shen and W. Yu, “Fractional Programming for Communication Systems—Part I: Power Control and Beamforming,” *IEEE Transactions on Signal Processing*, vol. 66, no. 10, pp. 2616–2630, 2018.
- [29] Z.-Q. Luo and T.-H. Chang, “SDP relaxation of homogeneous quadratic optimization: approximation bounds and applications,” in *Convex Optimization in Signal Processing and Communications*, D. P. Palomar and Y. C. Eldar, Eds. Cambridge University Press, 2009, pp. 117–165.
- [30] S. Boyd and L. Vandenberghe, *Convex Optimization*. USA: Cambridge University Press, 2004.
- [31] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization - Analysis, Algorithms, and Engineering Applications*. Philadelphia: SIAM, 2001.
- [32] L. Vandenberghe and S. Boyd, “Semidefinite Programming,” *SIAM Review*, vol. 38, no. 1, pp. 49–95, 1994.
- [33] Al-Kanj *et al.*, “Optimized joint cell planning and BS on/off switching for LTE networks,” *Wiley Wireless Communications and Mobile Computing*, vol. 16, no. 12, pp. 1537–1555, 2016.
- [34] Z. Dai *et al.*, “Nearly Optimal Sensor Selection for TDOA-Based Source Localization in Wireless Sensor Networks,” *IEEE Transactions on Vehicular Technology*, pp. 1–1, 2020.
- [35] L. Bedogni *et al.*, “The Bologna Ringway Dataset: Improving Road Network Conversion in SUMO and Validating Urban Mobility via Navigation Services,” *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5464–5476, 2015.
- [36] S. Hurley, “Planning effective cellular mobile radio networks,” *IEEE Transactions on Vehicular Technology*, vol. 51, no. 2, pp. 243–253, 2002.
- [37] E. Amaldi *et al.*, “Planning UMTS base station location: optimization models with power control and algorithms,” *IEEE Transactions on Wireless Communications*, vol. 2, no. 5, pp. 939–952, 2003.
- [38] L. Huang *et al.*, “Distributed coverage optimization for small cell clusters using game theory,” in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, 2013, pp. 2289–2293.
- [39] J. G. Andrews *et al.*, “A Tractable Approach to Coverage and Rate in Cellular Networks,” *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, 2011.
- [40] G. Koutitas *et al.*, “Deployment Strategies and Energy Efficiency of Cellular Networks,” *IEEE Transactions on Wireless Communications*, vol. 11, no. 7, pp. 2552–2563, 2012.
- [41] T. X. Brown, “Cellular performance bounds via shotgun cellular systems,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 11, pp. 2443–2455, 2000.
- [42] L. Chiaraviglio *et al.*, “Will the proliferation of 5G base stations increase the radio-frequency pollution?” in *IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, 2020.
- [43] J. T. J. Penttinen, “5G Network Planning and Optimization,” in *5G Explained: Security and Deployment of Advanced Mobile Communications*, 2019, pp. 255–269.
- [44] L. Chiaraviglio *et al.*, “Planning 5G Networks Under EMF Constraints: State of the Art and Vision,” *IEEE Access*, vol. 6, pp. 51 021–51 037, 2018.
- [45] A. Albanese, G. Encinas-Lago, V. Sciancalepore, X. Costa-Prez, D.-T. Phan-Huy, and S. Ros, “RIS-Aware Indoor Network Planning: The

- Rennes Railway Station Case,” in *ICC 2022 - 2022 IEEE International Conference on Communications (ICC)*, 2022, pp. 1–7.
- [46] J. G. Andrews *et al.*, “Modeling and Analyzing Millimeter Wave Cellular Systems,” *IEEE Trans. on Comm.*, vol. 65, no. 1, pp. 403–430, 2017.
- [47] A. L. Rezaabad *et al.*, “Ultra-Dense 5G Small Cell Deployment for Fiber and Wireless Backhaul-Aware Infrastructures,” *IEEE Transactions on Vehicular Technology*, vol. 67, no. 12, pp. 12 231–12 243, 2018.
- [48] D. Lopez-Perez *et al.*, “Towards 1 Gbps/UE in Cellular Systems: Understanding Ultra-Dense Small Cell Deployments,” *IEEE Communications Surveys Tutorials*, vol. 17, no. 4, pp. 2078–2101, 2015.
- [49] H. Wang *et al.*, “Coverage and Throughput Analysis with a Non-Uniform Small Cell Deployment,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 4, pp. 2047–2059, 2014.
- [50] A. Dong *et al.*, “A small cell deployment strategy towards amorphous coverage in the cellular network,” in *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2015, pp. 745–750.
- [51] K. C. Ho and M. Sun, “An accurate algebraic closed-form solution for energy-based source localization,” *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 8, pp. 2542–2550, 2007.
- [52] K. C. Ho and Wenwei Xu, “An accurate algebraic solution for moving source location using TDOA and FDOA measurements,” *IEEE Transactions on Signal Processing*, vol. 52, no. 9, pp. 2453–2463, 2004.
- [53] Yihong Qi *et al.*, “Analysis of wireless geolocation in a non-line-of-sight environment,” *IEEE Transactions on Wireless Communications*, vol. 5, no. 3, pp. 672–681, 2006.
- [54] Y. Han *et al.*, “Performance Limits and Geometric Properties of Array Localization,” *IEEE Transactions on Information Theory*, vol. 62, no. 2, pp. 1054–1075, 2016.
- [55] M. Z. Win *et al.*, “A Theoretical Foundation of Network Localization and Navigation,” *Proceedings of the IEEE*, vol. 106, no. 7, pp. 1136–1165, 2018.
- [56] A. Fascista *et al.*, “Millimeter-Wave Downlink Positioning With a Single-Antenna Receiver,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 9, pp. 4479–4490, 2019.
- [57] Z. Abu-Shaban *et al.*, “Error Bounds for Uplink and Downlink 3D Localization in 5G Millimeter Wave Systems,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 8, pp. 4939–4954, 2018.
- [58] F. Wen *et al.*, “5G Positioning and Mapping With Diffuse Multipath,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1164–1174, 2021.
- [59] H. Ko *et al.*, “GNSS Multipath-Resistant Cooperative Navigation in Urban Vehicular Networks,” *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5450–5463, 2015.
- [60] A. Nordio *et al.*, “Sensor Selection and Precoding Strategies for Wireless Sensor Networks,” *IEEE Transactions on Signal Processing*, vol. 63, no. 16, pp. 4411–4421, 2015.
- [61] N. Cao *et al.*, “Sensor Selection for Target Tracking in Wireless Sensor Networks With Uncertainty,” *IEEE Transactions on Signal Processing*, vol. 64, no. 20, pp. 5191–5204, 2016.
- [62] S. P. Chepuri and G. Leus, “Sparsity-Promoting Sensor Selection for Non-Linear Measurement Models,” *IEEE Transactions on Signal Processing*, vol. 63, no. 3, pp. 684–698, 2015.
- [63] X. Shen and P. K. Varshney, “Sensor Selection Based on Generalized Information Gain for Target Tracking in Large Sensor Networks,” *IEEE Transactions on Signal Processing*, vol. 62, no. 2, pp. 363–375, 2014.
- [64] A. Jindal and K. Psounis, “Modeling Spatially Correlated Data in Sensor Networks,” *ACM Transactions on Sensor Networks*, vol. 2, pp. 466–499, Nov. 2006.
- [65] S. Liu *et al.*, “Sensor Selection for Estimation with Correlated Measurement Noise,” *IEEE Transactions on Signal Processing*, vol. 64, no. 13, pp. 3509–3522, 2016.
- [66] M. Shamaiah *et al.*, “Greedy sensor selection: Leveraging submodularity,” in *49th IEEE Conference on Decision and Control (CDC)*, 2010, pp. 2572–2577.
- [67] Y. Zhao *et al.*, “Sensor Selection for TDOA-Based Localization in Wireless Sensor Networks With Non-Line-of-Sight Condition,” *IEEE Trans. on Vehicular Technology*, vol. 68, no. 10, pp. 9935–9950, 2019.



localization and prototyping.

Antonio Albanese (S'18) received the M.Sc. degree in Telecommunications Engineering from Politecnico di Milano, Italy, in 2018. Currently, he is pursuing his Ph.D. in Telematic Engineering at Universidad Carlos III de Madrid, Spain, while being appointed as Research Scientist at NEC Laboratories Europe GmbH, Heidelberg, Germany. His research field covers millimeter waves, reconfigurable intelligent surfaces, applied mathematical optimization and machine learning techniques, with a particular interest in



Transactions on Wireless Communications.

Vincenzo Sciancalepore (S'11–M'15–SM'19) received his M.Sc. degree in Telecommunications Engineering and Telematics Engineering in 2011 and 2012, respectively, whereas in 2015, he received a double Ph.D. degree. Currently, he is a Principal Researcher at NEC Laboratories Europe in Heidelberg, focusing his activity on network virtualization and network slicing challenges. He is the Chair of the ComSoc Emerging Technologies Initiative (ETI) on Reconfigurable Intelligent Surfaces (RIS) and an editor of *IEEE*



EPFL in 2015, 2013 and 2018 and a Fulbright scholar at University of Texas at Austin in 2019. Prof. Banchs authors over 150 publications in international conferences and journals, and is the co-inventor of several patents.

Albert Banchs (M'04–SM'12) received his M.Sc. and Ph.D. degrees from the Polytechnic University of Catalonia (UPC-BarcelonaTech) in 1997 and 2002, respectively. He is currently a Full Professor with the University Carlos III of Madrid (UC3M), with double affiliation as Deputy Director of the IMDEA Networks institute. Before joining UC3M, he was at ICSI Berkeley in 1997, at Telefonica I+D in 1998, and at NEC Europe Ltd. from 1998 to 2003. He was an Academic Guest at ETHZ in 2012, a Visiting Professor at



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 granted patents. Xavier received his Ph.D. degree in Telecommunications from the Polytechnic University of Catalonia (UPC) in Barcelona and was the recipient of a national award for his Ph.D. thesis.