

Radio over Fiber with Cascaded Structure: Algorithm for Uplink Positioning

Dexin Kong, *Graduate Student Member, IEEE*, Diana Pamela Moya Osorio, *Senior Member, IEEE*, and Erik G. Larsson, *Fellow, IEEE*

Abstract—Recent advancements in polymer microwave fiber (PMF) technology have created significant opportunities for robust, low-cost, and high-speed sub-terahertz (THz) radio-over-fiber communications. Recognizing these potential benefits, this paper explores a novel radio-over-fiber (RoF) structure that interconnects multiple radio units (RUs), booster units (BUs), and a central unit (CU) in cascade via fiber, envisioning its application in indoor scenarios. This structure creates a number of research opportunities when considering cascaded distortion effects introduced by non-linear power amplifiers (PAs) and the propagation channel over the fiber.

We propose maximum-likelihood and non-linear least-squares algorithms to estimate the entry RU and the time-of-arrival between the RoF and the user equipment, where estimating the entry RU is equivalent to estimating the propagation distance along the RoF stripe. For the case of linear PAs, we derive the Cramér-Rao lower bound to benchmark the performance of the estimators. Finally, we investigate the use of the system for uplink positioning. Our simulation results demonstrate that the proposed estimators perform satisfactorily even with the cascaded effects of non-linear PAs, and that the deployment of this RoF structure can enable new cost-effective opportunities for high-resolution positioning in indoor scenarios. In the numerical evaluation, we also use measured PMF characteristics for high-density polyethylene fibers.

Index Terms—Non-linear power amplifiers, polymer microwave fiber, radio over fiber, high-resolution positioning

I. INTRODUCTION

NEXT-generation wireless communication systems are expected to support ubiquitous wireless connectivity. To achieve this goal, one crucial limitation is the spectrum resources. Although sixth-generation (6G) communication systems are expected to adopt frequency bands that strike a balance between capacity and coverage, particularly within the new mid-band spectrum (7.125 to 24.25 GHz) [2], the sub-terahertz (sub-THz) range (90 to 300 GHz) is anticipated to enable high-precision sensing and short-range positioning applications [3]–[5]. Consequently, research on sub-THz technologies is expected to continue advancing [6]. In this work, we present a novel system operating in the sub-THz band to support ultra-high data rates, low latency, and enhanced

The authors are with the Dept. of Electrical Engineering (ISY), Linköping University, 581 83 Linköping, Sweden (email: dexin.kong@liu.se; diana.moya.osorio@liu.se; erik.g.larsson@liu.se).

This work is an extended version of our conference paper presented at IEEE SPAWC 2024 [1]. This work was supported in part by ELLIT, and in part by the 6G Tandem project funded by the European Union's Horizon Europe research and innovation programme under Grant Agreement No 101096302.

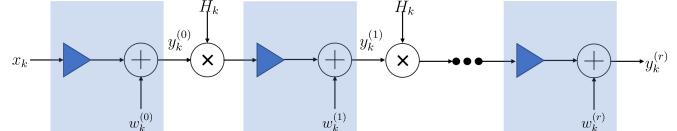


Fig. 1: Signal model for the RoF system with PAs and PMFs connected in cascade.

positioning accuracy. This is largely due to the wide bandwidth available in the sub-THz spectrum, which enables data transmission rates of several terabits per second and provides high spatial resolution.

Although there has been abundant research on the performance of sub-THz wireless communication systems [7]–[9], their implementation is still challenging [10]. Considering the cost of sub-THz radio hardware and the limited coverage [11], a low-cost and densely deployable solution is desirable. Recent developments in polymer microwave fiber (PMF) have motivated the development of radio-over-fiber (RoF) communications, which offer great opportunities for low-cost implementations of sub-THz wireless systems [12]–[14]. RoF systems can be implemented in different ways. Herein, we consider a system with a cascaded structure, where multiple radio units (RUs) and booster units (BUs) are interconnected via low-cost PMFs and finally connected to a central unit (CU). Each RU is equipped with antennas, a power amplifier (PA), and other radio frequency (RF) components; see Section III for details. The only difference between an RU and a BU is the absence of antennas in a BU. The RU capturing or transmitting the sub-THz signals acts as an access point, and all other units between the entry/exit RU and the CU only amplify and forward the signals. All signal processing is performed at the CU. Since there are no digital signal processing units in the RUs and BUs, the system is significantly less costly than a dense deployment of radio transceivers.

Fig. 1 illustrates the cascade RoF system under consideration (the different variables will be introduced later). In this system, the signal propagates over a short-range sub-THz wireless channel and over multiple dispersive fibers within the RoF system. The salient feature of this system is that the propagation distance over the sub-THz wireless channel is only a small fraction of the distance between the CU and the user equipment (UE). According to measurements reported in [15], state-of-the-art PMFs attenuate sub-THz signals less than 5 dB per meter in the D-band (110 to 170 GHz), while the

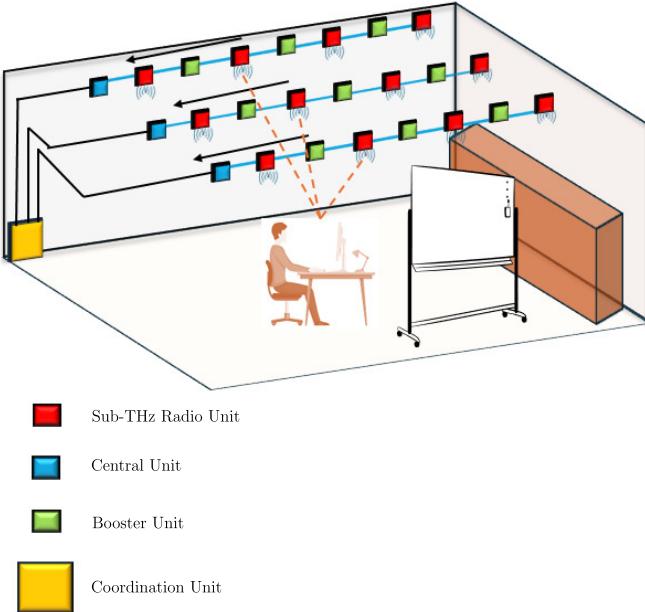


Fig. 2: An RoF system deployment in an indoor scenario.

free-space path loss in the D-band is around 70 dB on the first meter and then follows an exponential law [4]. Consequently, the overall propagation attenuation is significantly less in an RoF system compared to in a pure wireless sub-THz system with the same coverage.

The power amplifiers (PAs) are identified as the dominant source of power consumption. The energy efficiency (EE) of PAs is improved if a small back-off is adopted, i.e., the PAs are allowed to work in the non-linear regime [16]. Therefore, at each stage of the RoF system, the signals are distorted by a non-linear (PA) and a segment of dispersive PMF, with the latter being considered a linear time-invariant (LTI) system. After multiple stages, the signal will be subjected to a cumulative distortion, which we exploit to estimate the propagation distance over the RoF stripe. In this paper, each UE is assumed to transmit data to three RoF stripes that are deployed in parallel and mounted at the ceiling, as shown in Fig 2. The CU at each stripe jointly estimates the propagation distance over the stripe, as well as the time-of-arrival (ToA) with a clock bias between the UE and the stripe. Based on estimates of ToA and propagation distance from three CUs, trilateration is used to locate the UE. A detailed geometry model and problem formulation will be presented in Section IV.

Contributions: The specific contributions of this paper are (i) an analysis of the effects of non-linear PAs and dispersive PMFs connected in a cascade, and an analysis of the uplink (UL) signal; (ii) a maximum-likelihood (ML) framework and a non-linear least-squares (NLS) framework tailored to the cascade-structured RoF for estimating the entry RU and the ToA with bias; (iii) a thorough evaluation on the performance of the estimators with PAs working in the linear regime via a Cramér-Rao lower bound (CRLB) analysis; (iv) numerical results (using measured PMF dispersion characteristics and different values of non-linear factors) demonstrate that the proposed estimators perform well on the propagation distance

estimation even in the presence of the effects of cascaded non-linear PAs; and (v) we demonstrate the usability of the proposed system for high-resolution positioning in indoor scenarios. To the best of our knowledge, this is the first study of an RoF system with segments of dispersive PMFs and with non-linear amplifiers connected in cascade. The proposed system and analysis have the potential to open up several novel research directions and opportunities for positioning and localization systems.

The rest of this paper is organized as follows. Section II presents related works. Section III discusses the system overview and the structure of every component within an RoF stripe. Section IV formulates the problem and the corresponding signal models. Section V evaluates the ML estimator in the linear regime by the CRLB. Section VI shows the applied channel measurement data and numerical results. Section VI-C applies the proposed algorithm in an indoor positioning use case. Section VII concludes the paper.

Notations: Boldface letters denote column vectors; regular letters denote scalars. $(\cdot)^H$ denotes Hermitian transpose operation and $(\cdot)^T$ denotes the transpose operation; \mathbb{C} is the set of complex numbers; The k th element of a column vector \mathbf{x} is denoted by the subscript as x_k ; The superscript n of a vector/scalar $(\cdot)^{(n)}$ denotes the vector/scalar at the n th unit of an RoF stripe; $|\cdot|$ denotes the absolute value; $\|\cdot\|$ denotes the L^2 norm; $\delta(\cdot)$ is the delta function; $\det(\cdot)$ calculates the determinant of a matrix; $\text{diag}\{\cdot\}$ represents a diagonal matrix; $\mathbf{1}_N$ denotes a all one column vector with N elements; Finally, $j = \sqrt{-1}$ is the imaginary unit.

II. RELATED WORK

It is worth noting that this study considers an RoF system formed with only RF components and fibers. This implementation is novel and has not been investigated before, and it should not be confused with other RoF approaches that we introduce in the following.

In contrast to the system we consider, a traditional RoF system merges RF and optical fiber technologies, which typically contains lasers, an optical modulator/demodulator, optical amplifiers, optical fibers, a photomixer, and RF antennas. The mobile signal is carried by the modulated optical carriers and conveyed to the photomixer via a fiber. The photomixer then converts the optical carriers into RF electrical signals before the signals are transmitted by the RF antennas. Reference [17] discusses traditional RoF systems in more detail. It is broadly acknowledged that traditional RoF systems present low attenuation and high cost-efficiency due to the use of low-cost fibers whose development is reported in [12]–[14]. Moreover, since the signals in such RoF systems are well confined within the fiber, these systems are not restricted in bandwidth. The work in [18] experimentally demonstrates the RoF systems in the THz frequency band, which implements a fiber-wireless seamless network in an indoor scenario. The reported results demonstrate good performance in terms of carrier-to-noise ratio. However, this type of RoF system often suffers from severe non-linearity in optical components, such as lasers [19], and materials [20], which represent the main challenge for RoF communications.

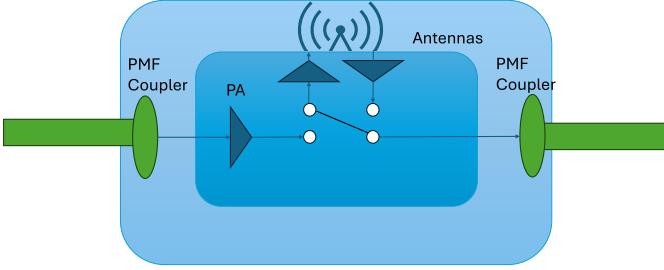


Fig. 3: The structure of an RU.

It is also worth differentiating the proposed system from radio stripes (RS), which is a commonly envisioned implementation of distributed massive MIMO (D-MMIMO). D-MMIMO distributes the antennas in a large area to provide ubiquitous service for every UE. The work in [21] provides a comprehensive overview of D-MMIMO in 6G. Although the D-MMIMO offers consistently high-quality service to all UEs, accurate synchronization between APs and the CU is challenging. On the other hand, RSs connect APs sequentially in one cable/fiber to the CU [22], which achieves simplified synchronization between APs via the shared bus. The authors in [23] investigate the high-accuracy positioning capability with RSs at the sub-6 GHz band. However, previous works about RS consider that every AP is equipped with full signal processing abilities, which are considerably expensive at THz frequencies. In contrast, the unique structure considered in this paper allows dense and flexible deployment because the CU allocates all the resources and processes the signals. The studied system distinguishes itself from the distributed massive MIMO in the literature due to the absence of multiple antenna signal processing.

III. SYSTEM OVERVIEW

Consider the system illustrated in Fig. 2, where multiple RoF stripes are deployed in parallel to serve UEs in an indoor scenario. Each RoF stripe consists of multiple RUs, BUs, and one CU at the end. One coordination unit connects the RoF stripes with the servers and coordinates different RoF stripes. All RoF stripes are installed on the ceiling at the same height. The configurations of RU, BU, CU, and PMF will be discussed below. All other components are considered to be LTI systems. In this study, we assume that those components introduce attenuation but no significant dispersion.

A. Polymer microwave fiber

For a unit-length fiber segment, the channel impulse response with L taps and amplitudes $([\beta_0, \beta_1, \dots, \beta_{L-1}]^T)$ ¹ can

¹This model is valid in a non-linear system if the signals are properly oversampled [24]. We consider a third-order polynomial non-linear distortion; hence, the signals must be oversampled at least three times.

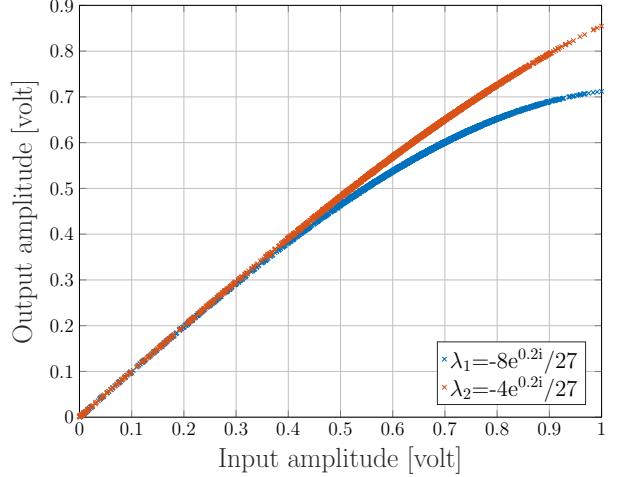


Fig. 4: Characteristics of the PA with different values of the non-linearity factor.

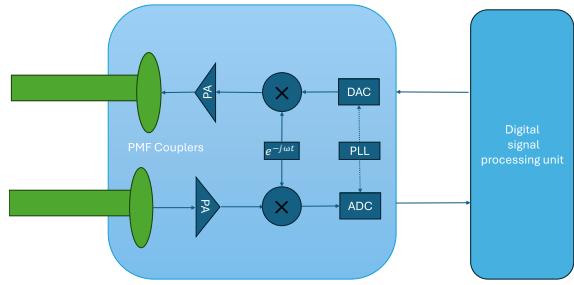


Fig. 5: The structure of a CU.

be expressed as

$$h_n = \sum_{l=0}^{L-1} \beta_l \delta(n-l), \quad (1)$$

where n is the time index. Then, a fiber with r unit length is equivalent to connecting r LTI systems with the same impulse response h_n in a cascade, whose impulse response is the r times convolution of h_n and is expressed as

$$\tilde{h}_n = \underbrace{h_n * h_n * \dots * h_n}_{r \text{ convolutions}}, \quad (2)$$

where all factors in (2) are modeled as in (1). In the frequency domain, we consider the frequency response for a set of K discrete frequencies, $\mathbf{f} = [f_0, \dots, f_{K-1}]^T$. Thus, the frequency domain representation of (2) is given by the multiplication of frequency responses. Let $\mathbf{H} = [H_0, \dots, H_{K-1}]^T$ represent the frequency response of a unit-length fiber. Then, the frequency response of a unit-length fiber at the k th frequency H_k is the Fourier transform of its impulse response given by

$$H_k = \sum_{n=0}^{N-1} h_n e^{-j2\pi f_k n T_s}, \quad (3)$$

where T_s is the sampling time interval. Hence, the frequency response of an r -units-long fiber, at the k th frequency, can be expressed as

$$\tilde{H}_k = \underbrace{H_k H_k \cdots H_k}_{r \text{ multiplications}} = H_k^r. \quad (4)$$

B. Radio Unit & Booster Unit

RUs are responsible for capturing incoming sub-THz signals, amplifying them, and forwarding them to the next component in the same stripe. Each RU consists of PMF couplers, PAs, antennas, and a switch as illustrated in Fig. 3. For the operation of the RoF system, we consider that each RU is equipped with a switch that selectively receives signals, either from the PMF or the sub-THz antennas. When the RU detects signals from the antennas, the switch disconnects the PMF input, allowing only the signals captured by the antennas to pass through. Therefore, each RU only admits signals either from the antennas or from the PMF. If there are two RUs in a stripe capturing signals, only the RU that is closer to the CU would act as an access point.

In case the RU structure does not implement a switch, thus allowing to receive signals coming from both sides, antennas and the PMF, a superposition of the received signals from different RUs would be received at the CU. This case is not addressed in the present work, as its treatment requires an entirely separate analysis.

BUs are deployed between any two adjacent RUs in an RoF stripe, as shown in Fig. 2. As previously mentioned, the BUs have the same structure as the RUs without the presence of antennas. The BUs amplify the signals and forward them to the next component.

To understand the impact of PAs, we examine both regimes of operation of PAs, linear and non-linear. The effects of non-linear PAs are characterized with a memoryless polynomial model. We use a third-order memoryless polynomial model with a factor λ , as higher-order terms are significantly less than the third-order term [25], [26]. A real-valued non-linear factor indicates only amplitude distortion, while a complex-valued non-linear factor indicates both phase and amplitude distortion. In practice, for a specific PA, λ can be measured in a laboratory. The relationship between the input x_n and output y_n of a PA is given by

- Linear regime: $y_n = Gx_n$,
- Non-linear regime: $y_n = G(x_n + \lambda x_n |x_n|^2)$,

where G is the amplification coefficient. Fig. 4 illustrates the characteristics of PA with different values of non-linear factors, where we can observe that the non-linearity arises with a signal magnitude of ~ 0.4 volts.

Due to the absence of a digital signal processing unit at a RU/BU, every RU/BU is incapable of processing signals. This design lowers the price of each RU/BU, but also disables the RU/BU to mitigate signal impairments introduced by PAs and PMFs. As a consequence, this results in the accumulation of signal distortion along the propagation in an RoF system.

C. Central unit

The CU is in charge of allocating resources, synchronizing the units in a stripe, and processing data. Each CU comprises PAs, mixers, couplers, ADC/DAC, phase-locked loops (PLL), and a digital signal processing unit (see Fig. 5). The CU is capable to operate in full-duplex mode to communicate with one UE simultaneously.

IV. PROBLEM FORMULATION AND SIGNAL MODEL

To localize a UE in a two-dimensional space, at a minimum, three observations are required. With the system structure of the cascaded-structured RoF, three stripes are required to receive the signals from a UE simultaneously. As shown in Fig. 6, each stripe jointly estimates the propagation distance (denoted as r) in the stripe as well as the ToA (denoted as τ) of the incoming signals. The estimated propagation distances are used to locate three RUs that serve the UE, and the estimated ToAs are used for trilateration.

Before deriving the received signal model, we emphasize the key assumptions of the paper. The RoF stripes are assumed to be synchronized in time and phase via wired connections. From the perspective of practical operation, it is challenging to synchronize the UE and RoF stripes. Therefore, it is assumed that the UE has a clock bias and phase offset relative to all RoF stripes, which are unknown to the CU. The sub-THz signals propagate both over the air and over the fiber. Due to the sparse nature of the sub-THz wireless channel, the line-of-sight (LoS) components are considered to be dominant, whose path loss and ToA are not available to the CU. On the other hand, a fixed-length fiber segment is an LTI system whose impulse response can be measured in a laboratory. According to (1) to (4), the channel information of a fiber of any length is available to the CU. Similarly, the PA characteristics (G , λ , and noise variance at room temperature σ^2) are stable in the room temperature and can be measured after manufacture. In practice, a small deviation of PA characteristics does not have a significant effect, which is discussed in Section IV-D. In conclusion, the unknown parameters are i) A , a complex value absorbing the effects of the wireless channel and the phase offset between the UE and the RoF; ii) τ , the propagation delay including ToA and a clock bias between the UE and the stripe; iii) r , the propagation distance in a stripe; iv) the coordinates of the UE. In particular, each CU/stripe estimates one propagation delay and propagation distance in the stripe, which are then used for trilateration to obtain the coordinates of the UE. All other variables in the signal model, as shown in the sections below, are known.

The RUs and BUs are uniformly distributed, which means that the distance between two consecutive units on an RoF stripe is fixed, and the distance between two adjacent RoF stripes is also fixed. Note that the stripes are shifted slightly relative to one another to ensure that the UE is not served by three RUs that lie on a line. This ensures that the geometric dilution of precision (GDOP) is favorable (see calculation in Section VI-C). The coordinate of the i th position of UE is given by $(U_x^{(i)}, U_y^{(i)}, U_z)$, and the location of the CU in the

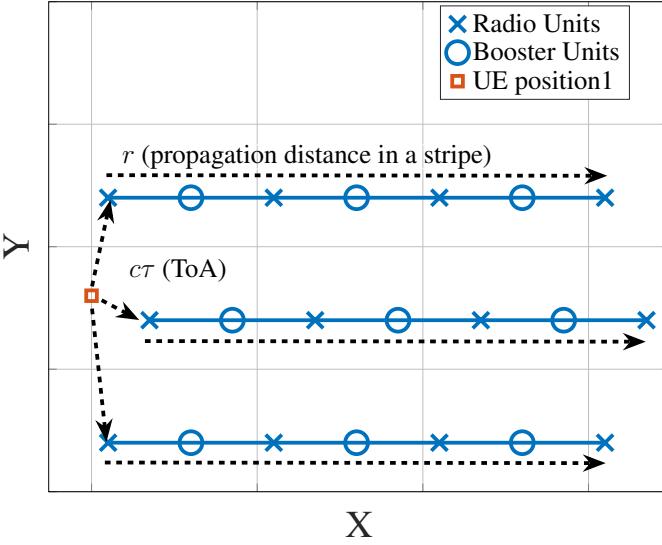


Fig. 6: Positioning of the UE with three RoF stripes.

m th stripe is given by $(P_x^{(m)}, P_y^{(m)}, 0)$. Therefore, the distance between the UE and the RU in service at the m th stripe is

$$d_{m,i} = \sqrt{|U_x^{(i)} - (P_x^{(m)} - r)|^2 + |U_y^{(i)} - P_y^{(m)}|^2 + U_z^2}. \quad (5)$$

The distance between the UE and the ceiling, U_z , is assumed to be fixed and known to the CU. The proposed positioning algorithm is based on ToA with bias, whose GDOP of a certain UE position $(U_x^{(i)}, U_y^{(i)})$ can be calculated according to the standard formulas [27]. The GDOP is defined as

$$GDOP(i) = \sqrt{\text{tr} \left(\mathbf{H}_{\text{ToA}}^T(i) \mathbf{H}_{\text{ToA}}(i) \right)^{-1}}, \quad (6)$$

where $\mathbf{H}_{\text{ToA}}(i)$ is obtained from (5) and can be written as

$$\mathbf{H}_{\text{ToA}}(i) = \begin{bmatrix} \frac{U_x^{(i)} - P_x^{(1)} + r}{d_{1,i}}, \frac{U_y^{(i)} - P_y^{(1)} + r}{d_{1,i}} \\ \frac{U_x^{(i)} - P_x^{(2)} + r}{d_{2,i}}, \frac{U_y^{(i)} - P_y^{(2)} + r}{d_{2,i}} \\ \frac{U_x^{(i)} - P_x^{(3)} + r}{d_{3,i}}, \frac{U_y^{(i)} - P_y^{(3)} + r}{d_{3,i}} \end{bmatrix}. \quad (7)$$

It should be noted that this paper emerges as an initial study of new fundamental problems that appear when LTI systems and non-linear systems are connected in cascade. Depending on how the system is implemented and in what way the switches are configurable, r might be known as a priori, in which case its estimation would be redundant. The initial working assumption of the architecture, however, was that there may not be a centrally controllable switch, such that the CU relies on the cumulative signal distortion of the cascaded LTI/non-linear propagation structure to estimate r .

A. RU with PAs operating in the linear regime

Following the discussion in Section III, this paper considers PAs working in both linear and non-linear regimes. In this subsection, let us consider PAs working in the linear regime. Under this scenario, the whole RoF system is LTI. Thus, the signal model considers the dispersive channel of PMFs and the wireless channel from the UE to the entry RU. It is assumed

that the UE transmits a known sequence of samples, $\mathbf{s} = [s_0, \dots, s_{K-1}]^T \in \mathbb{C}^{K \times 1}$ over K frequency points. Initially, the signals undergo the wireless channel between the UE and the RU, which is unknown to the CU.

Thus, the k th element of the input signal at the RoF, x_k with $\mathbf{x} = [x_0, x_1, \dots, x_{K-1}]^T$ (see Fig. 1), can be written as

$$x_k = A e^{-j2\pi f_k \tau} s_k, \quad (8)$$

where τ can be expressed as

$$\tau = \frac{d}{c} + \delta_t, \quad (9)$$

with c being the speed of light, and δ_t being the clock offset between the UE and the RoF. Note that distinguishing the effects of d and δ_t is not possible since they have the same influence over the signals, i.e., phase shift and delay (see (9)). However, one can rely on the effects of cascaded fibers and PAs to estimate r and τ separately.

The k th element of the output signal at the entry RU, denoted as $y_k^{(0)}$, is given by

$$y_k^{(0)} = G A e^{-j2\pi f_k \tau} s_k + w_k^{(0)}, \quad (10)$$

with $w_k^{(r)} \sim \mathcal{CN}(0, \sigma^2)$ being the k th AWGN noise component at the r th stage. The wireless channel as such does not contribute noise; only the PAs add noise, as they are the only active electronic components.

In a general form, after propagating through r units (i.e., r PMF segments propagation distance) along the RoF, the received signal at the CU, with $\mathbf{y}^{(r)} = [y_0^{(r)}, y_1^{(r)}, \dots, y_{K-1}^{(r)}]^T$, can be written as

$$y_k^{(r)} = G^{r+1} H_k^r A e^{-j2\pi f_k \tau} s_k + \underbrace{G^r H_k^r w_k^{(0)} + \dots + w_k^{(r)}}_{(r+1) \text{ independent noises}}. \quad (11)$$

Note that the CU also amplifies the signals. Hence, there are $(r+1)$ PAs and r PMF segments.

The cumulative filtered noise by linear PAs and fibers still follows the Gaussian distribution since the linear transformation does not change the probability distribution. Let \tilde{w}_k denote the sum of independent Gaussian noise terms, then the variance of \tilde{w}_k is given by

$$\text{Var}(\tilde{w}_k) = \sum_{r'=0}^r b_k^{r'} \sigma^2. \quad (12)$$

where $b_k \triangleq (G|H_k|)^2$. Until now, the parameter r takes discrete values, indicating the number of stages that the sub-THz signals propagate in an RoF stripe. Since CRLB only applied to continuous-valued variables, we propose a continuous-value representation of r in Proposition 1.

Proposition 1: *The discrete-valued r in (12) can be approximated as a continuous value by splitting one RU into infinitely dense virtual RUs with length $\Delta r \approx 0$. The response of one virtual RU is $(GH_k)^{\Delta r}$. By accepting this approximation, the response of one RU is $[(GH_k)^{\Delta r}]^{\frac{1}{\Delta r}}$, which is equivalent to GH_k . The variance of effective noise at (12) can be rewritten as*

$$\text{Var}(\tilde{w}_k) = \sigma^2 \frac{\ln b_k}{b_k - 1} \int_0^{r+1} (b_k)^{r'} dr'. \quad (13)$$

Proof: the variance of \tilde{w}_k in (12) can be re-expressed as the sum of a geometric series as

$$\sum_{r'=0}^r b_k^{r'} \sigma^2 = \frac{1 - b_k^{r+1}}{1 - b_k} \sigma^2, \quad (14)$$

which is equal to the result of (13) as

$$\begin{aligned} \sigma^2 \frac{\ln b_k}{b_k - 1} \int_0^{r+1} (b_k)^{r'} dr' &= \frac{\ln b_k}{b_k - 1} \frac{b_k^{r+1} - 1}{\ln b_k} \sigma^2, \\ &= \frac{b_k^{r+1} - 1}{b_k - 1} \sigma^2. \end{aligned} \quad (15)$$

Note that (15) only holds for $b_k \neq 1, \forall k$. In other words, there are two cases worth being studied. We name them as *flat fiber* and *frequency-selective fiber* cases. This approach was also adopted in [28] to calculate the CRLB of a discrete variable.

1) *Flat fiber*: Let us consider a special case in which the frequency response of the fiber is totally flat, that is, the case for a system with very limited bandwidth. With the use of PMF, the frequency response can be very flat at 1GHz bandwidth. It is assumed that the amplification G is set to compensate for the attenuation, corresponding to $b_k = 1, \forall k$. Consequently, the variance of effective noise is

$$\Sigma = (r + 1)\sigma^2 \mathbf{I}_K, \quad (16)$$

and the phase shift induced by the fiber grows linearly with the propagation distance r . In other words, a flat fiber only supports one-mode transmission within the fiber, corresponding to a single-tap impulse response. It should be noted that the system can not separate the effects of the wireless channel and the RoF based on the phase, but only relies on the noise variance shown in (16) to estimate r .

The unknown parameters to be estimated in this problem are $\theta = [A, r, \tau]$, of which the parameters r and τ lead to a non-linear problem while A is involved linearly. We employ a maximum likelihood (ML) estimator to tackle this estimation problem. First, we define the vector $\mathbf{g} \in \mathbb{C}^{K \times 1}$ as

$$\mathbf{g} = G^{r+1} [e^{-j2\pi f_0 \tau} H_0^r s_0, \dots, e^{-j2\pi f_{K-1} \tau} H_{K-1}^r s_{K-1}]^T. \quad (17)$$

The log-likelihood function can be written as

$$\ln P(\mathbf{y}^{(r)} | \theta) = \ln(\pi \det(\Sigma)) + (\mathbf{y}^{(r)} - A\mathbf{g})^H \Sigma^{-1} (\mathbf{y}^{(r)} - A\mathbf{g}). \quad (18)$$

Since A is linearly involved in the likelihood function and irrelevant to Σ , the estimation of A can be obtained as

$$\hat{A} = \min_A \|\mathbf{y}^{(r)} - A\mathbf{g}\|^2. \quad (19)$$

It is not difficult to obtain its estimate as

$$\hat{A} = \frac{\mathbf{g}^H \mathbf{y}^{(r)}}{\|\mathbf{g}\|^2}. \quad (20)$$

After dropping the dependency on A by utilizing \hat{A} and combining with the expression of Σ in (16), we can rewrite the objective function as

$$\begin{bmatrix} \hat{r} \\ \hat{\tau} \end{bmatrix} = \min_{r, \tau} \ln((r+1)K\pi\sigma^2) + \frac{\|\mathbf{y}^{(r)} - \mathbf{g}(\mathbf{g}^H \mathbf{g})^{-1} \mathbf{g}^H \mathbf{y}^{(r)}\|^2}{(r+1)\sigma^2}. \quad (21)$$

Therefore, after a two-dimensional grid search, the estimate \hat{r} as well as $\hat{\tau}$, can be obtained. The estimate of A is calculated as (20) by plugging in the estimates \hat{r} and $\hat{\tau}$.

2) *Frequency-selective fiber*: A frequency-selective fiber means the magnitude of its frequency response fluctuates over the bandwidth of interest, which makes b_k not equal to one. Without the loss of generality, we assume $b_k \neq 1, \forall k$, which ensures the validity of (15) for all frequencies. This is the case for a system with a large bandwidth, for example, 10GHz. Consequently, the variance of effective noise is

$$\Sigma = \text{diag} \left\{ \frac{b_0^{r+1} - 1}{b_0 - 1} \sigma^2, \frac{b_1^{r+1} - 1}{b_1 - 1} \sigma^2, \dots, \frac{b_{K-1}^{r+1} - 1}{b_{K-1} - 1} \sigma^2 \right\}, \quad (22)$$

and the phase shifts introduced by the fiber might grow non-linearly with the propagation distance r . Similarly, one can view a *frequency-selective fiber* to support multi-mode transmission, whose corresponding impulse response has multiple taps.

With the noise covariance being (22), prewhitening is applied by filtering the received signal $\mathbf{y}^{(r)}$ and the vector \mathbf{g} , which can be written as

$$\begin{aligned} \tilde{\mathbf{y}}^{(r)} &= \Sigma^{-\frac{1}{2}} \mathbf{y}^{(r)}, \\ \tilde{\mathbf{g}} &= \Sigma^{-\frac{1}{2}} \mathbf{g}. \end{aligned} \quad (23)$$

Then, the estimate of A is obtained by (20), and the estimates of r and τ are calculated as (21).

B. RU with PAs operating in the non-linear regime

For EE purposes, the PAs are allowed to work in the non-linear regime, as modeled in Section III. As previously stated, the UE transmits a known sequence of samples $\mathbf{s} = [s_0, \dots, s_{K-1}]^T$ over K frequency points to the RoF. We model the input signal in the time domain as

$$x_n = \frac{1}{K} \sum_{k=0}^{K-1} e^{j2\pi \frac{n}{N} k} A e^{-j2\pi f_k \tau} s_k, \quad (24)$$

with $x_n \in [x_0, x_1, \dots, x_{K-1}]^T$.

The output of the entry RU can be written as

$$y_n^{(0)} = G \left(x_n + \underbrace{\lambda x_n |x_n|^2}_{\text{PA's non-linearity}} \right) + w_n^{(0)}. \quad (25)$$

To obtain the output at the following units, we define the function $f(\cdot)$, which combines the effects of a non-linear PA and one segment of PMF, and can be expressed as

$$f(y_n^{(0)}) = G \left(\sum_{l=0}^{L-1} \beta_l y_{n-l}^{(0)} + \lambda \sum_{l=1}^{L-1} \beta_l y_{n-l}^{(0)} \left| \sum_{l=1}^{L-1} \beta_l y_{n-l}^{(0)} \right|^2 \right). \quad (26)$$

If $f(\cdot)$ in (26) takes a vector as input, it operates element-wise.

The signal $\mathbf{y}^{(0)}$ would undergo r PAs and r segments of PMF to reach the CU. In other words, the signal $\mathbf{y}^{(0)}$ would undergo a recursive process r times. Let $f^r(\mathbf{y}^{(0)})$ be the overall function, which can be expressed as

$$f^r(\mathbf{y}^{(0)}) = \underbrace{f(f(f(\dots(\mathbf{y}^{(0)}))))}_{r \text{ times}}. \quad (27)$$

Therefore, the received signal at the CU can be expressed as $\mathbf{y}^{(r)} = f^r(\mathbf{y}^{(0)}) + \mathbf{w}^{(r)}$, where all the unknown parameters are interwoven with each other. As a result, we cannot apply the same estimator proposed in the linear regime ((20) and (21)). Alternatively, we can employ an NLS framework.

With PAs operating in the non-linear regime, an NLS framework may not serve as an optimal estimator as the noise distribution is transformed from a Gaussian distribution into another probability distribution. However, the output of an NLS framework will not deviate significantly from the optimum. Following the NLS framework, the objective function w.r.t. the unknown parameter A, r , and τ is given by

$$\begin{bmatrix} \hat{A} \\ \hat{\tau} \\ \hat{r} \end{bmatrix} = \min_{A, \tau, r} \left\| \mathbf{y}^{(r)} - f^r \left(G \left(x_n + \lambda x_n |x_n|^2 \right) \right) \right\|^2, \quad (28)$$

where a three-dimensional grid search is needed to obtain the estimates. To solve this problem, we employ a particle swarm optimization (PSO) algorithm described in Algorithm 1, to reduce the complexity compared to a three-dimensional grid search.

Compared to the ML estimator in (18), the employed NLS framework in the non-linear regime disregards the information contained in the noise covariance.

C. Positioning of the UE

According to (5), the distance between the i th UE and the entry RUs is written as $\mathbf{d} \triangleq [d_{1,i}, d_{2,i}, d_{3,i}]^T$. The ToA from the UE to the entry RUs is interwoven with a clock offset as defined in (9), which can be viewed as a bias. The position of the UE $U_x^{(i)}, U_y^{(i)}$ is estimated through the estimated $\tau = (1/c)\mathbf{d} + \mathbf{1}_3\delta_t$ at three RoF stripes and can be written as

$$\begin{bmatrix} \hat{U}_x^{(i)} \\ \hat{U}_y^{(i)} \end{bmatrix} = \min_{U_x^{(i)}, U_y^{(i)}, \delta_t} \left\| \boldsymbol{\tau} - \frac{\mathbf{d}}{c} - \mathbf{1}_3\delta_t \right\|^2, \quad (29)$$

where the estimation of δ_t is a linear model with the rest of quantities. Let $\mathbf{R} \triangleq \boldsymbol{\tau} - \frac{\mathbf{d}}{c}$. It is not difficult to write the estimate of δ_t as

$$\hat{\delta}_t = \frac{\mathbf{1}_3^T \mathbf{R}}{\mathbf{1}_3^T \mathbf{1}_3}. \quad (30)$$

Plugging in $\hat{\delta}_t$ back to (29), the estimate of the UE positions can be written as

$$\begin{bmatrix} \hat{U}_x^{(i)} \\ \hat{U}_y^{(i)} \end{bmatrix} = \min_{U_x^{(i)}, U_y^{(i)}} \left\| \mathbf{R} - \mathbf{1}_3(\mathbf{1}_3^T \mathbf{1}_3)^{-1} \mathbf{1}_3^T \mathbf{R} \right\|^2. \quad (31)$$

D. The effect of the non-ideal PA

The analysis of this section is based on ideal PA assumptions, i.e., all PAs have identical and constant characteristics. Although this assumption is ideal, a small deviation from it does not lead to significant effects on the estimators. Practically, the noise variance and PA gain show little variation with the room temperature (much less than 2 dB gain variation in [29]). The cumulative impact of cascading PAs is much more dominant than that of varying individual PA characteristics. For example, two PAs in a cascade doubles the noise variance and squares the gain, compared to a single PA.

V. CRAMÉR-RAO LOWER BOUND

CRLB is a very useful tool for evaluating the performance of any unbiased estimator, as it provides a measure of the lower limit of the variance. However, the CRLB for the case of PAs operating in the non-linear regime leads to a cumbersome problem due to the presence of iterative non-linearity in the signal model. Alternatively, this section introduces the CRLB of the proposed ML estimator with PAs operating in the linear regime as a benchmark of estimation accuracy.

For simplicity, we express the complex valued A and H with their magnitudes $|A|$ and $|H_k|$ and phases ϕ and ψ_k , i.e., $A = |A|e^{j\phi}$ and $H = |H_k|e^{j\psi_k}$. Hence, the unknown parameters are defined as $\boldsymbol{\theta} \triangleq [|A|, \phi, \tau, r]^T$. The definition for the CRLB is given by

$$[\text{Var}(\hat{\boldsymbol{\theta}}_i)] \geq [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{i,i}, \quad (32)$$

where $\mathbf{I}(\boldsymbol{\theta})$ is the Fisher information matrix (FIM) for the vector $\boldsymbol{\theta}$ and is defined as

$$\begin{aligned} [\mathbf{I}(\boldsymbol{\theta})]_{i,j} &= \text{tr} \left\{ \mathbf{C}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_i} \mathbf{C}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial \theta_j} \right\} \\ &+ 2\Re \left\{ \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})^H}{\partial \theta_i} \mathbf{C}(\boldsymbol{\theta})^{-1} \frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \theta_j} \right\}, \end{aligned} \quad (33)$$

which reveals the amount of information associated with each pair of unknown parameters. The FIM in (33) is valid only for Gaussian models. One important characteristic of FIM is that every FIM is positive semi-definite, indicating the information content is non-negative and ensuring its eigenvalues are non-negative. Moreover, a positive definite FIM indicates the parameters are identifiable, and a singular FIM indicates the parameters are non-identifiable.

As mentioned in Section IV, there are two expressions for the variance of effective noise as shown in (16) and in (22). Naturally, the CRLB needs to be calculated for both flat fiber and frequency-selective fiber cases. Detailed calculations can be found in the Appendix A.

A. CRLB for flat fiber

Now, let us start with the flat fiber case, whose $\mathbf{C}(\boldsymbol{\theta})$ and $\boldsymbol{\mu}(\boldsymbol{\theta})$ are defined as

$$\begin{aligned} \mathbf{C}(\boldsymbol{\theta}) &= (r+1)\sigma^2 \mathbf{I}_K \\ \boldsymbol{\mu}(\boldsymbol{\theta}) &= \left[G^{r+1} |A| |H_0|^r s_0 e^{j(\phi+r\psi_0-2\pi f_0 \tau)}, \dots, \right. \\ &\quad \left. G^{r+1} |A| |H_{K-1}|^r s_{K-1} e^{j(\phi+r\psi_{K-1}-2\pi f_{K-1} \tau)} \right]^T. \end{aligned} \quad (34)$$

According to the symmetry property, $[\mathbf{I}(\boldsymbol{\theta})]_{i,j} = [\mathbf{I}(\boldsymbol{\theta})]_{j,i}$ is always valid so that there are 10 unique entries in each FIM. Regarding the *flat fiber*, the elements of the FIM are

$$[\mathbf{I}(\boldsymbol{\theta})]_{|A|,|A|} = 2 \frac{G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2, \quad (35a)$$

$$[\mathbf{I}(\boldsymbol{\theta})]_{|A|,\phi} = 0, \quad (35b)$$

$$[\mathbf{I}(\boldsymbol{\theta})]_{|A|,\tau} = 0, \quad (35c)$$

$$[\mathbf{I}(\theta)]_{|A|,r} = 0, \quad (35d)$$

$$[\mathbf{I}(\theta)]_{\phi,\phi} = 2 \frac{G^2 |A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2, \quad (35e)$$

$$[\mathbf{I}(\theta)]_{\phi,\tau} = - \frac{4\pi G^2 |A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k |s_k|^2, \quad (35f)$$

$$[\mathbf{I}(\theta)]_{\phi,r} = 2 \frac{G^2 |A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 \psi_k, \quad (35g)$$

$$[\mathbf{I}(\theta)]_{\tau,\tau} = \frac{8\pi^2 G^2 |A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k^2 |s_k|^2, \quad (35h)$$

$$[\mathbf{I}(\theta)]_{r,r} = \frac{K}{(r+1)^2} + 2 \frac{G^2 |A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 \psi_k^2, \quad (35i)$$

$$[\mathbf{I}(\theta)]_{\tau,r} = - \frac{4\pi G^2 |A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 f_k \psi_k. \quad (35j)$$

Discussions: The main takeaways from the CRLB calculation, from (35a) to (35j), are:

- The estimate of $|A|$ is solely dependent on the received signal-to-noise ratio (see (35a)).
- As shown in (35f), (35g), and (35j), there is a coupling between the estimates of τ , r , and ϕ , which is proved to make the FIM close to a singular matrix through numerical results in the following section.
- The phase shifts introduced by the fiber are interwoven with the phase shifts introduced from the wireless channel and the phase offset, making it impossible for the system to identify the propagation distance r based on the phase information.
- According to (35i), the system can rely on the overall noise variance to estimate the propagation distance, as it cumulates linearly with r .

B. CRLB for frequency-selective fiber

For the frequency-selective fiber case, the noise variance is defined as

$$\begin{aligned} \mathbf{C}(\boldsymbol{\theta}) &= \text{diag} \left(\sigma^2 \frac{\ln b_0}{b_0 - 1} \int_0^{r+1} (b_0)^{r'} dr', \dots, \right. \\ &\quad \left. \sigma^2 \frac{\ln b_{K-1}}{b_{K-1} - 1} \int_0^{r+1} (b_{K-1})^{r'} dr' \right), \\ [\mathbf{C}^{-1}(\boldsymbol{\theta})]_{k,k} &= \frac{b_k - 1}{\sigma^2 (b_k^{r+1} - 1)}, \\ \left[\frac{\partial \mathbf{C}(\boldsymbol{\theta})}{\partial r} \right]_{k,k} &= \frac{\sigma^2 b_k^{r+1} \ln b_k}{b_k - 1}, \end{aligned} \quad (36)$$

and $\boldsymbol{\mu}(\boldsymbol{\theta})$ is the same as (34). Similarly, the entries in the FIM can be obtained as

$$[\mathbf{I}(\theta)]_{|A|,|A|} = 2G^2 \sum_{k=0}^{K-1} \frac{(b_k - 1)b_k^r |s_k|^2}{\sigma^2 (b_k^{r+1} - 1)}, \quad (37a)$$

$$[\mathbf{I}(\theta)]_{|A|,\phi} = 0, \quad (37b)$$

$$[\mathbf{I}(\theta)]_{|A|,\tau} = 0, \quad (37c)$$

$$[\mathbf{I}(\theta)]_{|A|,r} = 2G^2 |A| \sum_{k=0}^{K-1} \frac{(b_k - 1)b_k^r |s_k|^2 \ln \sqrt{b_k}}{\sigma^2 (b_k^{r+1} - 1)}, \quad (37d)$$

$$[\mathbf{I}(\theta)]_{\phi,\phi} = 2G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k - 1)b_k^r |s_k|^2}{\sigma^2 (b_k^{r+1} - 1)}, \quad (37e)$$

$$[\mathbf{I}(\theta)]_{\phi,\tau} = - 4\pi G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k - 1)f_k b_k^r |s_k|^2}{\sigma^2 (b_k^{r+1} - 1)}, \quad (37f)$$

$$[\mathbf{I}(\theta)]_{\phi,r} = 2G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k - 1)b_k^r |s_k|^2 \psi_k}{\sigma^2 (b_k^{r+1} - 1)}, \quad (37g)$$

$$[\mathbf{I}(\theta)]_{\tau,\tau} = 8\pi^2 G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k - 1)f_k^2 b_k^r |s_k|^2}{\sigma^2 (b_k^{r+1} - 1)}, \quad (37h)$$

$$[\mathbf{I}(\theta)]_{\tau,r} = - 4\pi G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k - 1)b_k^r |s_k|^2 f_k \psi_k}{\sigma^2 (b_k^{r+1} - 1)}, \quad (37i)$$

$$\begin{aligned} [\mathbf{I}(\theta)]_{r,r} &= 2G^2 |A|^2 \sum_{k=0}^{K-1} \frac{|s_k|^2 (b_k - 1)b_k^r ((\ln \sqrt{b_k})^2 + \psi_k^2)}{\sigma^2 (b_k^{r+1} - 1)} \\ &\quad + \sum_{k=0}^{K-1} \left(\frac{b_k^r \ln b_k}{1 - b_k^{r+1}} \right)^2. \end{aligned} \quad (37j)$$

Discussion: Based on the CRLB obtained for the frequency-selective fiber, one can observe

- With the use of frequency-selective fiber, a coupling between $|A|$ and r appears as illustrated in (37d), which comes from the varied magnitude of the frequency response of the PMF.
- As illustrated in (37j), the system can rely on the shape of the spectrum of the received signal at CU, while it can only rely on the noise variance with a flat fiber (see (35i)).
- The coupling between ϕ , τ , and r still exists (see (37g), (37f), and (37i)).

VI. SIMULATION RESULTS

In this section, we first validate the proposed NLS estimator, relying on the measured channel of a segment of PMF. Then, we validate the proposed CRLB with two channels generated under a constraint on the total energy. Moreover, we apply the proposed positioning algorithm in an indoor scenario.

A. Validation of the proposed estimators with measured fiber channel characteristics

In this section, we illustrate the performance obtained by the proposed estimators with the fiber channel measurements presented in [15]. The channel measurement data in [15] was obtained by transmitting sub-THz (D-band) signals over a one-meter long high-density polyethylene PMF with a solid rectangular cross-section of 1 mm \times 2 mm and a density of 0.93 g/cm³. These measurements inherently include several imperfections, such as noise, interference, and measurement errors. Therefore, a median filtering with a window size of 300 is implemented through MATLAB to smooth the data in group delay and magnitude, separately. Both the smoothed data and the original data are shown in Figs. 7a and 7b.

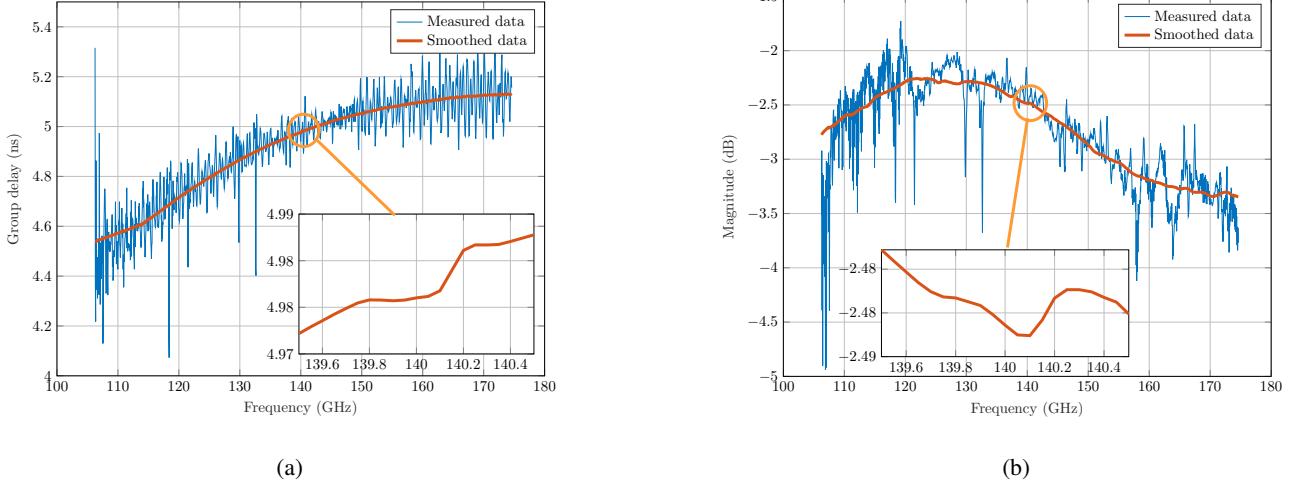


Fig. 7: Measured characteristics of a 1-meter PMF made of high-density polyethylene with a solid rectangular cross-section in the D-band (110 GHz to 170 GHz) [15]: (a) Group delay; (b) Magnitude.

The measured data consists of channel characteristics at K discrete frequencies, encompassing both magnitudes ($|\mathbf{H}| = [|H_0|, \dots, |H_{K-1}|]^T$) and group delay. Subsequently, the phase response $\psi = [\psi_0, \dots, \psi_{K-1}]^T$ was obtained by integrating the group delay on the discrete frequencies. Then, the transfer function can be expressed as

$$H_k = |\mathbf{H}_k| \exp(j\psi_k). \quad (38)$$

The employed LoS model comprises free-space path loss and a large-scale fading component, which can be written as

$$|A| = G_t G_r \left(\frac{\lambda}{2\pi d} \right)^2 + \zeta, \quad (39)$$

with λ being the wavelength, d being defined in (5), G_t and G_r being the transmitter gain, and receiver gain respectively. The large-scale fading component ζ follows a log-normal distribution with a standard deviation of 2 dB. This model can be replaced by other developed LoS models for sub-THz propagation, such as the 3GPP model in [30]. In our simulation, G_t is 36 dB, and G_r is also 36 dB. Assuming the use of state-of-the-art patch antennas with 9 dBi gain [31], [32], a 22×22 array is required, which translates to $4 \text{ cm} \times 4 \text{ cm}$ antenna size. A horn antenna with equivalent gain can also be used.

Monte Carlo simulations were performed with an RoF system comprising five RUs and BUs, and one CU. Between every two units, a fiber with the frequency response as expressed in (38) serves as the connection. Distortion incurred by PAs was illustrated in Fig. 4. The UE transmits a known sequence of samples using a quadrature phase shift keying waveform with a unit amplitude. The UE accesses the RoF system through the third RU. The channel response shown in Fig. 7b corresponds to a spectrum with 1 GHz bandwidth centered around 140 GHz, from which the magnitude in a one-meter fiber is approximately -2.48 dB. Therefore, G is set to 2.48 dB for PAs. The parameters in (39) were set to ensure the input amplitude is in a certain regime (linear or non-linear) of PAs.

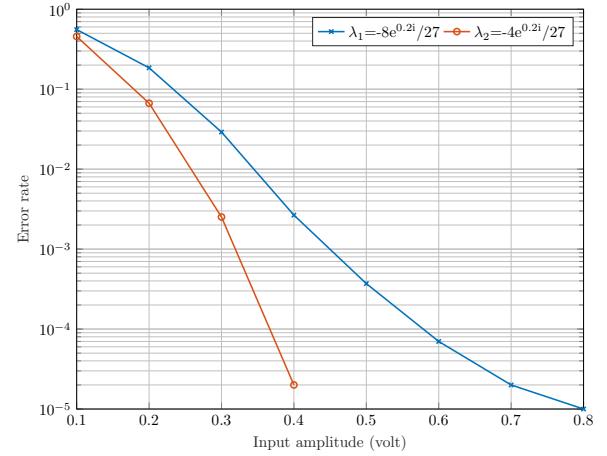


Fig. 8: The error rate of $\hat{r} \neq r$.

The proposed NLS estimator was implemented in Matlab, and the results are shown in Fig. 8. It is observed that the error rate would decrease to a level of 10^{-5} with a proper input amplitude, which indicates that the proposed algorithm is able to accurately estimate the propagation distance even when PAs work in the non-linear regime. Through comparison between the estimation accuracy with two non-linear factors, one can observe an accuracy loss with a higher non-linear factor.

B. Validation of CRLB with assumed fiber channels

To validate the CRLB calculated, extensive Monte-Carlo simulations were performed and based on two artificial channels shown in Fig. 9, corresponding to a flat fiber (red curve) and a frequency-selective fiber (blue curve). Each channel contains the same amount of energy, i.e., $\sum_{k=0}^{K-1} |H_k|^2 = \mathcal{E}$. It should be emphasized that the shape of the spectrum of a frequency-selective fiber does not impose any deviation in

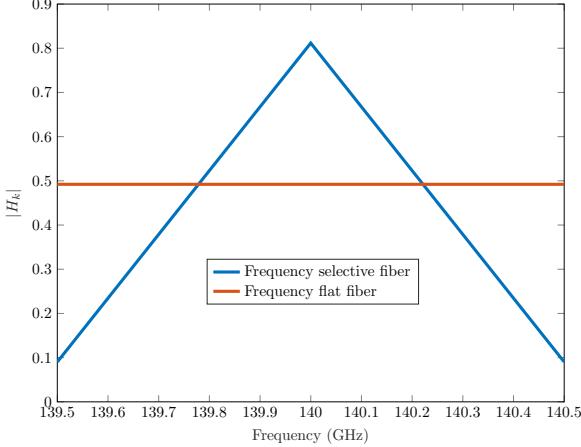


Fig. 9: The magnitude of the frequency response for the frequency flat fiber and the frequency-selective fibers.

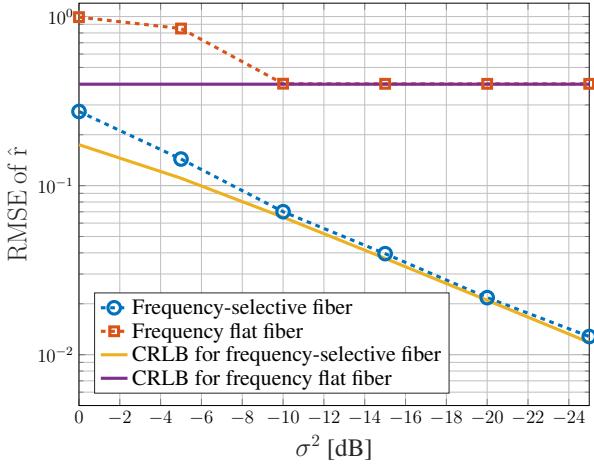


Fig. 10: RMSE and CRLB of \hat{r} , with frequency flat and frequency-selective fibers.

the conclusions of this paper. The corresponding root mean squared error (RMSE) of target parameters was calculated and plotted in Figs. 10, and 11.

The comparison between the CRLB w.r.t. \hat{r} with a flat fiber and that with a non-flat fiber indicates that the estimation accuracy is enhanced with a frequency-selective fiber, which proves the fact that the system can extract additional information from the spectrum of the received signal. Another observation from Fig. 10 is a flat bound with a flat fiber. This understanding can be acquired from (35i), where the CU estimates the noise variance and determines its quantitative relationship with the noise variance σ^2 at each PA. Therefore, this bound is independent of the exact value of σ^2 . The ML estimator achieves an RMSE equal to the CRLB, indicating that it attains optimal performance. Similar conclusions can be drawn from Fig 11. It is worth noting that the gap between the RMSE of certain estimates and the corresponding CRLB does not necessarily indicate suboptimality, as the simulation grids are inherently constrained.

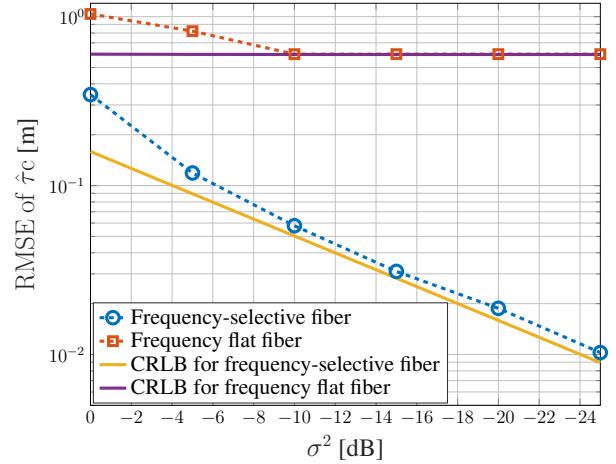


Fig. 11: RMSE and CRLB of $\hat{\tau}$, with frequency flat and frequency-selective fibers.

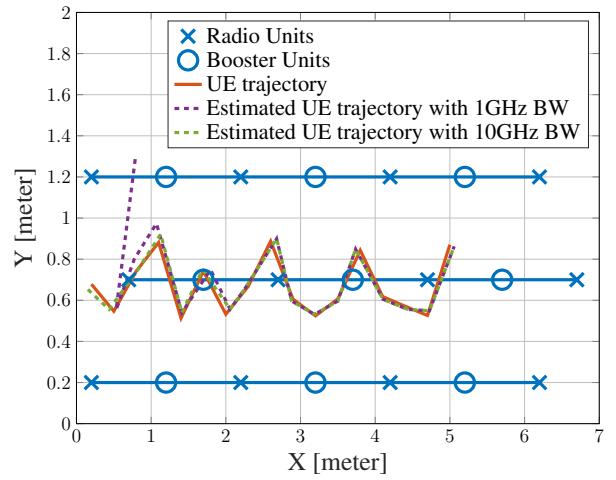
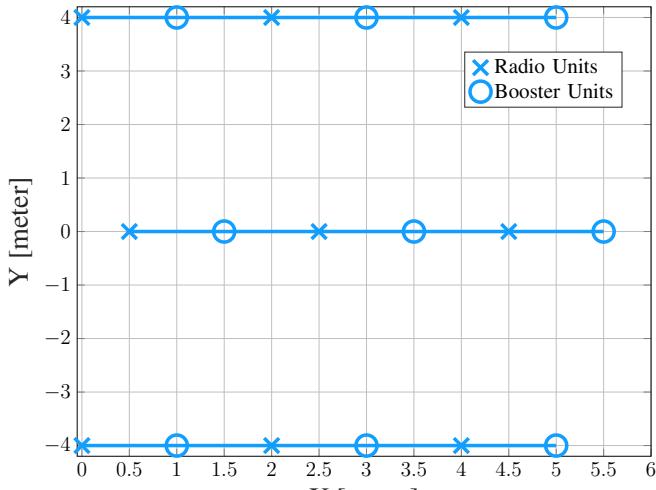


Fig. 12: Floorplan of an indoor scenario with three RoF stripes, a UE trajectory, and the estimated UE trajectory with 1 GHz and 10 GHz bandwidth.

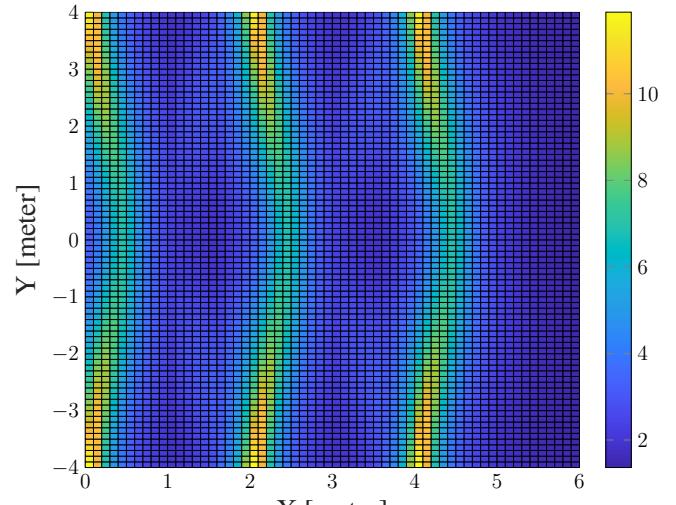
C. Use case: Indoor positioning

A pertinent use case for the proposed estimation framework on the considered RoF system is indoor positioning. This low-cost design allows the operator to densely deploy RoF stripes for extreme data rate and good coverage. Here, we assume a UE moves along a trajectory in an indoor scenario illustrated in Fig. 2, where three RoF stripes are serving the UE simultaneously. To ensure fairness and avoid potential biases, no additional assumptions about the antennas are made. Instead, it is simply assumed that the UE is consistently served by three RUs located at three RoF stripes. The height of the UE is assumed to be consistently 1.5 meters below the ceiling at all positions, with a uniform spacing of 1 meter between adjacent units in stripes and a spacing of 0.5 meters between adjacent RoF stripes.

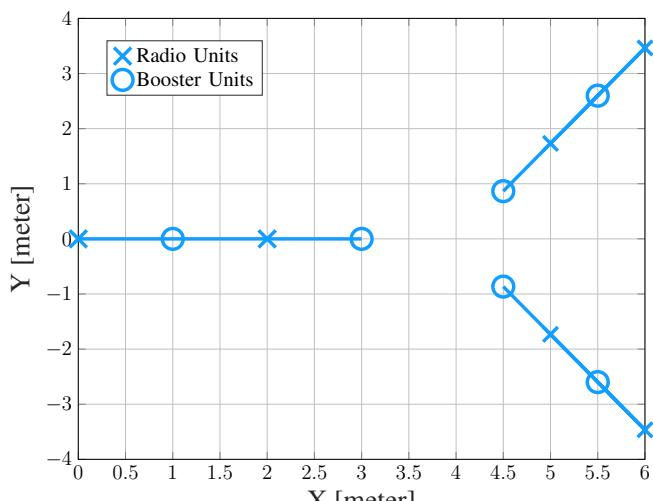
The simulations were carried out using bandwidths of 1 GHz and 10 GHz, corresponding to delay resolutions of 0.3 meters and 0.03 meters, respectively. Fig 12 shows the estimated UE trajectories with different bandwidths. The RMSE of the



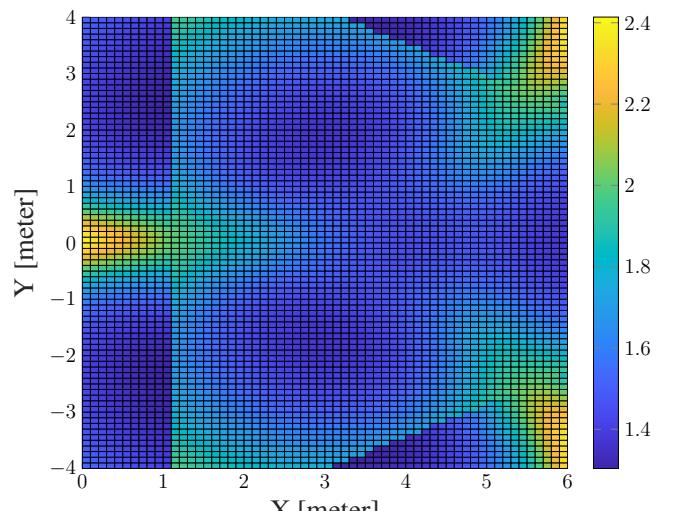
(a)



(b)



(c)



(d)

Fig. 13: Comparison of RU deployments in parallel and star topologies: (a) parallel-topology network, (b) GDOP distribution for the parallel topology, (c) star-topology network, and (d) GDOP distribution for the star topology.

estimated positions, obtained through Monte Carlo simulations, is 0.04 meters for 10 GHz and 0.5 meters for 1 GHz, both closely approaching their respective delay resolutions. The GDOP quantifies the influence of the geometry on the positioning performance. The proposed positioning algorithm is based on ToA with bias, whose GDOP can be calculated according to (5)-(6). Stripe deployments can follow different network topologies, such as the parallel topology, which is the focus of this paper, or a star topology (see Figs. 13a and 13c). For consistency, both topologies are deployed to cover the same service area; however, one should notice that the star topology deploys fewer RUs than the parallel topology. Figs. 13b and 13d illustrate the GDOP distribution for both topologies. In Figs. 13b and 13d, the high GDOP in certain positions is mainly attributed to the alignment of two RUs along a line, which results in an unfavorable geometry

and reduced positioning performance. Additionally, the star network topology exhibits greater geometric diversity than the parallel network topology. A topic for future research can be to investigate optimal RU deployments by jointly considering the positioning and communication performance.

VII. CONCLUSION

This paper studied a novel, low-cost, and easily deployable sub-THz RoF system implementation that combines PMFs and RUs in a cascaded structure. A UL signal model of this system was developed by considering two regimes for PAs, linear and non-linear. For our analysis, we considered two types of fibers: flat fibers and frequency-selective fibers, whose effects were included in the signal model. Based on this model, an ML framework and an NLS framework were developed to estimate the propagation distance along the RoF and the TOA of signals

Algorithm 1: Particle Swarm Optimization Algorithm

```

1 Function Cost( $\theta, \lambda, G, s, f(x), \mathbf{y}^{(r)}$ ):
2    $x_n = \frac{1}{K} \sum_{k=0}^{K-1} e^{j2\pi\frac{n}{N}k} A e^{-j2\pi f_k \tau} s_k;$ 
3    $c \leftarrow \left\| \mathbf{y}^{(r)} - f^r (G(x_n + \lambda x_n |x_n|^2)) \right\|^2$ 
4   return  $c$ ;
5 Function PSO( $\lambda, G, H, s, \mathbf{y}^{(r)}, f(x)$ ):
6    $m \leftarrow 50$ ; // Set the #iterations
7    $p \leftarrow 100$ ; // Set particle size
8    $w_1 \leftarrow 1.5$ ; // personal best pos. Weights
9    $w_2 \leftarrow 1.5$ ; // global best pos. Weights
10   $w_{\text{ine}} \leftarrow 0.3$ ; // Set inertia weight
11   $n \leftarrow 4$ ; // #variables
12   $c_{\text{gbest}} \leftarrow 0$ ;  $v \leftarrow 0$ ; // Initialize global cost
13   $\theta_{\text{min}} \leftarrow [|A|_{\text{min}}, \phi_{\text{min}}, \tau_{\text{min}}, r_{\text{min}}];$ 
14   $\theta_{\text{max}} \leftarrow [|A|_{\text{max}}, \phi_{\text{max}}, \tau_{\text{max}}, r_{\text{max}}];$ 
15  for  $i \leftarrow 1$  to  $p$  do
16    for  $j \leftarrow 1$  to  $n$  do
17       $\theta(i, j) \leftarrow (\theta_{\text{max}}(j) - \theta_{\text{min}}(j)) * \text{rand}(1) + \theta_{\text{min}}(j);$ 
18      // Set the initial values
19       $\theta_{\text{aux}}(j) \leftarrow \theta(i, j);$ 
20       $c(i) \leftarrow \text{Cost}(\theta_{\text{aux}}, \lambda, G, s, f(x), \mathbf{y}^{(r)});$ 
21      if  $c(i) < c_{\text{gbest}}$  then
22         $c_{\text{gbest}} \leftarrow c(i);$ 
23        for  $j \leftarrow 1$  to  $n$  do
24           $\theta_{\text{gbest}}(j) \leftarrow \theta(i, j);$  // Iterate the
25          // global best cost
26   $\theta_{\text{pbest}} \leftarrow \theta$ ;  $c_{\text{pbest}} \leftarrow c;$ 
27  for  $i \leftarrow 1$  to  $m$  do
28     $v \leftarrow w_{\text{ine}} * v + w_1 * \text{rand}(p, n) * (\theta_{\text{pbest}} - \theta) + w_2 *$ 
29     $\text{rand}(p, n) * (\theta_{\text{gbest}} - \theta);$  // Move particles
30     $\theta \leftarrow \theta + v;$ 
31    for  $j \leftarrow 1$  to  $p$  do
32      for  $k \leftarrow 1$  to  $n$  do
33        if  $\theta(j, k) < \theta_{\text{min}}(k)$  then
34           $\theta(j, k) \leftarrow \theta_{\text{min}}(k);$ 
35        if  $\theta(j, k) > \theta_{\text{max}}(k)$  then
36           $\theta(j, k) \leftarrow \theta_{\text{max}}(k);$ 
37         $\theta_{\text{aux}}(k) \leftarrow \theta(j, k);$ 
38         $c(j) \leftarrow \text{Cost}(\theta_{\text{aux}}, \lambda, G, s, f(x), \mathbf{y}^{(r)});$ 
39        if  $c(j) < c_{\text{pbest}}(j)$  then
40           $c_{\text{pbest}}(j) \leftarrow c(j);$ 
41          for  $k \leftarrow 1$  to  $n$  do
42             $\theta_{\text{pbest}}(j, k) \leftarrow \theta(j, k);$ 
43          if  $c(j) < c_{\text{gbest}}$  then
44             $c_{\text{gbest}} \leftarrow c(j);$ 
45            for  $k \leftarrow 1$  to  $n$  do
46               $\theta_{\text{gbest}}(k) \leftarrow \theta(j, k);$ 
47             $w_{\text{ine}} \leftarrow w_{\text{ine}} * 0.7;$ 
48  return  $c_{\text{gbest}}, \theta_{\text{gbest}};$ 

```

transmitted from a UE. The CRLB was derived as a benchmark of the variance of the proposed ML estimator.

Monte-Carlo simulations were performed to assess the performance of the proposed estimator. Our results demonstrate that good performance of the proposed estimators can be achieved under the effects of cascaded PMFs and RUs. The ML estimator achieves superior performance for a frequency-selective fiber compared to a frequency-flat fiber. Finally, our proposed algorithms show high accuracy in positioning UEs in an indoor scenario.

A future direction could be estimating how the second moment of the variance is affected by the non-linearities and including that in the estimation criterion. Another possible direction is to replace the third-order polynomial model of PA non-linearity with a more accurate one. Also, considering the antenna radiation patterns, such as a highly directional antenna, and handling the UE's orientation, enhances the research.

ACKNOWLEDGMENT

We thank Dr. Frida Strömbeck and Prof. Herbert Zirath from Chalmers University of Technology for generously sharing the measured PMF characteristics that we used in our numerical examples.

APPENDIX

In this appendix, we provide the detailed derivations of the CRLB shown in Section V. Specifically, we write the respective derivations of the CRLB for both the flat-fiber and frequency-selective fiber assumptions. Note that both CRLBs are obtained from (12). The two types of fiber show different characteristics. Hence, we present the CRLB for each case separately. We define $\Omega_k \triangleq \phi + r\psi_k - 2\pi f_k \tau$.

CRLB FOR FLAT FIBER

Partial derivatives in (33) can be calculated as

$$\left[\frac{\partial \mu(\theta)}{\partial |A|} \right]_k = G e^{j(\Omega_k)} s_k, \quad (40a)$$

$$\left[\frac{\partial \mu(\theta)}{\partial \phi} \right]_k = jG|A|s_k e^{j(\Omega_k)}, \quad (40b)$$

$$\left[\frac{\partial \mu(\theta)}{\partial \tau} \right]_k = -j2\pi f_k G|A|s_k e^{j(\Omega_k)}, \quad (40c)$$

$$\left[\frac{\partial \mu(\theta)}{\partial r} \right]_k = G|A|e^{j(\Omega_k)} s_k (j\psi_k). \quad (40d)$$

The derivations of CRLB are as follows.

$$\begin{aligned} [\mathbf{I}(\theta)]_{|A|, |A|} &= 2\Re \left\{ \frac{G^{(2r+2)}}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |H_k|^{2r} |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)} \right\}, \\ &= 2 \frac{G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2, \end{aligned} \quad (41a)$$

$$\begin{aligned} [\mathbf{I}(\theta)]_{|A|, \phi} &= 2\Re \left\{ \frac{j|A|G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)} \right\}, \\ &= 0, \end{aligned} \quad (41b)$$

$$\begin{aligned} [\mathbf{I}(\theta)]_{|A|, \tau} &= 2\Re \left\{ -\frac{j2\pi G^2 |A|}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)} \right\}, \\ &= 0, \end{aligned} \quad (41c)$$

$$[\mathbf{I}(\theta)]_{|A|,r} = 2\Re \left\{ \frac{G^2|A|}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 (j\psi_k) \right\}, \\ = 0, \quad (41\text{d})$$

$$[\mathbf{I}(\theta)]_{\phi,\phi} = 2\Re \left\{ \frac{-j^2|A|^2G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)} \right\}, \\ = 2 \frac{G^2|A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2, \quad (41\text{e})$$

$$[\mathbf{I}(\theta)]_{\phi,\tau} = 2\Re \left\{ \frac{j^2 2\pi|A|^2G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)} \right\}, \\ = - \frac{4\pi G^2|A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k |s_k|^2, \quad (41\text{f})$$

$$[\mathbf{I}(\theta)]_{\phi,r} = 2\Re \left\{ \frac{-j|A|^2G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 (j\psi_k) \right\}, \\ = 2 \frac{G^2|A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 \psi_k, \quad (41\text{g})$$

$$[\mathbf{I}(\theta)]_{\tau,\tau} = 2\Re \left\{ \frac{-j^2 4\pi^2|A|^2G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k^2 |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)} \right\}, \\ = \frac{8\pi^2 G^2|A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k^2 |s_k|^2, \quad (41\text{h})$$

$$[\mathbf{I}(\theta)]_{r,r} = K \left\{ \frac{\sigma^2}{(r+1)\sigma^2} \frac{\sigma^2}{(r+1)\sigma^2} \right\} \\ + 2\Re \left\{ \frac{|A|^2G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 (-j\psi_k) (j\psi_k) \right\}, \\ = \frac{K}{(r+1)^2} + 2 \frac{G^2|A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 \psi_k^2, \quad (41\text{i})$$

$$[\mathbf{I}(\theta)]_{\tau,r} = 2\Re \left\{ \frac{j2\pi|A|^2G^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} f_k |s_k|^2 (j\psi_k) \right\}, \\ = - \frac{4\pi G^2|A|^2}{(r+1)\sigma^2} \sum_{k=0}^{K-1} |s_k|^2 f_k \psi_k. \quad (41\text{j})$$

CRLB FOR FREQUENCY-SELECTIVE FIBER

One can derive the first-order derivatives in (36) as

$$\left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial |A|} \right]_k = G^{(r+1)} |H_k|^r e^{j(\phi+r\psi_k-2\pi f_k \tau)} s_k, \quad (42\text{a})$$

$$\left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \phi} \right]_k = jG^{(r+1)} |A| |H_k|^r s_k e^{j(\phi+r\psi_k-2\pi f_k \tau)}, \quad (42\text{b})$$

$$\left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial \tau} \right]_k = -j2\pi f_k G^{(r+1)} |A| |H_k|^r s_k e^{j(\phi+r\psi_k-2\pi f_k \tau)}, \quad (42\text{c})$$

$$\left[\frac{\partial \boldsymbol{\mu}(\boldsymbol{\theta})}{\partial r} \right]_k = G^{(r+1)} |A| |H_k|^r s_k e^{j(\phi+r\psi_k-2\pi f_k \tau)} \\ \times (\ln \sqrt{b_k} + j\psi_k). \quad (42\text{d})$$

The derivations of CRLB as follows.

$$[\mathbf{I}(\theta)]_{|A|,|A|} = 2\Re \left\{ G^2 \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)}}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = 2G^2 \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2}{\sigma^2(b_k^{r+1}-1)}, \quad (43\text{a})$$

$$[\mathbf{I}(\theta)]_{|A|,\phi} = 2\Re \left\{ j|A| G^2 \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)}}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = 0, \quad (43\text{b})$$

$$[\mathbf{I}(\theta)]_{|A|,\tau} = 2\Re \left\{ -j2\pi G^2 |A| \sum_{k=0}^{K-1} \frac{(b_k-1)f_k b_k^r |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)}}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = 0, \quad (43\text{c})$$

$$[\mathbf{I}(\theta)]_{|A|,r} = 2\Re \left\{ G^2 |A| \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2 (\ln |GH_k| + j\psi_k)}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = 2G^2 |A| \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2 \ln \sqrt{b_k}}{\sigma^2(b_k^{r+1}-1)}, \quad (43\text{d})$$

$$[\mathbf{I}(\theta)]_{\phi,\phi} = 2\Re \left\{ |A|^2 G^2 \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)}}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = 2G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2}{\sigma^2(b_k^{r+1}-1)}, \quad (43\text{e})$$

$$[\mathbf{I}(\theta)]_{\phi,\tau} = 2\Re \left\{ -2\pi |A|^2 G^2 \sum_{k=0}^{K-1} \frac{(b_k-1)f_k b_k^r |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)}}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = -4\pi G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k-1)f_k b_k^r |s_k|^2}{\sigma^2(b_k^{r+1}-1)}, \quad (43\text{f})$$

$$[\mathbf{I}(\theta)]_{\phi,r} = 2\Re \left\{ -j|A|^2 G^2 \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2 (\ln |GH_k| + j\psi_k)}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = 2G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k-1)b_k^r |s_k|^2 \psi_k}{\sigma^2(b_k^{r+1}-1)}, \quad (43\text{g})$$

$$[\mathbf{I}(\theta)]_{\tau,\tau} = 2\Re \left\{ 4\pi^2 |A|^2 G^2 \sum_{k=0}^{K-1} \frac{(b_k-1)f_k^2 b_k^r |s_k|^2 e^{-j(\Omega_k)} e^{j(\Omega_k)}}{\sigma^2(b_k^{r+1}-1)} \right\}, \\ = 8\pi^2 G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k-1)f_k^2 b_k^r |s_k|^2}{\sigma^2(b_k^{r+1}-1)}, \quad (43\text{h})$$

$$\begin{aligned}
[I(\theta)]_{r,r} &= \text{tr} \left\{ C(\theta)^{-1} \frac{\partial C(\theta)}{\partial r} C(\theta)^{-1} \frac{\partial C(\theta)}{\partial r} \right\} + 2\Re \left\{ |A|^2 G^2 \sum_{k=0}^{K-1} \frac{(b_k - 1)|s_k|^2 b_k^r (\ln |GH_k| - j\psi_k) (\ln |GH_k| + j\psi_k)}{\sigma^2(b_k^{r+1} - 1)} \right\}, \\
&= \sum_{k=0}^{K-1} \left(\frac{b_k^r \ln b_k}{1 - b_k^{r+1}} \right)^2 + 2G^2 |A|^2 \sum_{k=0}^{K-1} \frac{|s_k|^2 (b_k - 1) b_k^r ((\ln \sqrt{b_k})^2 + \psi_k^2)}{\sigma^2(b_k^{r+1} - 1)},
\end{aligned} \tag{43i}$$

$$\begin{aligned}
[I(\theta)]_{\tau,r} &= 2\Re \left\{ j2\pi |A|^2 G^2 \sum_{k=0}^{K-1} \frac{(b_k - 1) f_k b_k^r |s_k|^2 (\ln |GH_k| + j\psi_k)}{\sigma^2(b_k^{r+1} - 1)} \right\}, \\
&= -4\pi G^2 |A|^2 \sum_{k=0}^{K-1} \frac{(b_k - 1) b_k^r |s_k|^2 f_k \psi_k}{\sigma^2(b_k^{r+1} - 1)}.
\end{aligned} \tag{43j}$$

REFERENCES

[1] D. Kong, D. P. Moya Osorio, and E. G. Larsson, "Propagation distance estimation for radio over fiber with cascaded structure," in *Proc. of 2024 IEEE 25th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2024, pp. 691–695.

[2] A. Bazzi, R. Bomfin, M. Mezzavilla, S. Rangan, T. Rappaport, and M. Chaffi, "Upper mid-band spectrum for 6G: Vision, opportunity and challenges," *arXiv preprint arXiv:2502.17914*, 2025.

[3] Z. Chen, X. Ma, B. Zhang, Y. Zhang, Z. Niu, N. Kuang, W. Chen, L. Li, and S. Li, "A survey on Terahertz communications," *China Communications*, vol. 16, no. 2, pp. 1–35, 2019.

[4] X. Cai, X. Cheng, and F. Tufvesson, "Toward 6G with Terahertz communications: Understanding the propagation channels," *IEEE Communications Magazine*, vol. 62, no. 2, pp. 32–38, 2024.

[5] A. Bourdoux, A. N. Barreto, B. van Liempd, C. de Lima, D. Dardari, D. Belot, E.-S. Lohan, G. Seco-Granados, H. Sarieddeen, H. Wymeersch *et al.*, "6G white paper on localization and sensing," *arXiv preprint arXiv:2006.01779*, 2020.

[6] M. Shafi, E. G. Larsson, X. Lin, D. Panaitopol, S. Parkvall, F. Ronteix-Jacquet, and A. Toskala, "Industrial viewpoints on RAN technologies for 6G," *arXiv preprint arXiv:2508.08225*, 2025.

[7] H. Elayan, O. Amin, R. M. Shubair, and M.-S. Alouini, "Terahertz communication: The opportunities of wireless technology beyond 5G," in *Proc. of 2018 International Conference on Advanced Communication Technologies and Networking (CommNet)*, 2018, pp. 1–5.

[8] C. Chaccour, W. Saad, O. Semiai, M. Bennis, and P. Popovski, "Joint sensing and communication for situational awareness in wireless THz systems," in *Proc. of IEEE International Conference on Communications (ICC)*, 2022, pp. 3772–3777.

[9] O. Kanhere and T. S. Rappaport, "Outdoor sub-THz position location and tracking using field measurements at 142 GHz," in *Proc. of IEEE International Conference on Communications (ICC)*, 2021, pp. 1–6.

[10] P. Heydari, "Terahertz integrated circuits and systems for high-speed wireless communications: Challenges and design perspectives," *IEEE Open Journal of the Solid-State Circuits Society*, vol. 1, pp. 18–36, 2021.

[11] E. Björnson, L. Van der Perre, S. Buzzi, and E. G. Larsson, "Massive MIMO in sub-6 GHz and mmWave: Physical, practical, and use-case differences," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 100–108, 2019.

[12] A. Standaert, M. Rousstia, S. Sinaga, and P. Reynaert, "Analysis and experimental verification of the HE11 mode in hollow PTFE fibers," in *Proc. of 2015 Asia-Pacific Microwave Conference (APMC)*, vol. 3, 2015, pp. 1–3.

[13] M. De Wit, Y. Zhang, and P. Reynaert, "Analysis and design of a foam-cladded PMF link with phase tuning in 28-nm CMOS," *IEEE Journal of solid-state circuits*, vol. 54, no. 7, pp. 1960–1969, 2019.

[14] P. Reynaert, M. Tytgat, W. Volckaerts, A. Standaert, Y. Zhang, M. De Wit, and N. Van Thienen, "Polymer microwave fibers: a blend of RF, copper and optical communication," in *Proc. of ESSCIRC Conference 2016: 42nd European Solid-State Circuits Conference*, 2016, pp. 15–20.

[15] F. Strömbeck, "Integrated Circuit Design for High Data Rate Polymer Microwave Fiber Communication," 2023, Doctoral thesis. [Online]. Available: <https://research.chalmers.se/en/publication/534754>

[16] M. Sarajlić, N. Tervo, A. Pärssinen, L. H. Nguyen, H. Halbauer, K. Roth, V. Kumar, T. Svensson, A. Nimir, S. Zeitz, M. Dörpinghaus, and G. Fettweis, "Waveforms for sub-THz 6G: Design guidelines," in *Proc. of 2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*, 2023, pp. 168–173.

[17] J. Beas, G. Castanon, I. Aldaya, A. Aragon-Zavala, and G. Campuzano, "Millimeter-wave frequency radio over fiber systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 4, pp. 1593–1619, 2013.

[18] M. Sung, S. Kim, E.-S. Kim, S.-R. Moon, M. Kim, I.-M. Lee, K. H. Park, J. Ki Lee, and S.-H. Cho, "Design of RoF-based fiber-wireless system for THz-band 6G indoor network," in *Proc. of 2022 European Conference on Optical Communication (ECOC)*, 2022, pp. 1–4.

[19] J. Riepl, J. Raab, P. Abajyan, H. Nong, J. Freeman, L. Li, E. H. Linfield, A. G. Davies, A. Wacker, T. Albes *et al.*, "Field-resolved high-order sub-cycle nonlinearities in a Terahertz semiconductor laser," *Light: Science & Applications*, vol. 10, no. 1, p. 246, 2021.

[20] L. Jia, J. Wu, Y. Zhang, Y. Qu, B. Jia, and D. J. Moss, "Third-order optical nonlinearities of 2D materials at telecommunications wavelengths," *Micromachines*, vol. 14, no. 2, p. 307, 2023.

[21] H. Q. Ngo, G. Interdonato, E. G. Larsson, G. Caire, and J. G. Andrews, "Ultradense Cell-Free Massive MIMO for 6G: Technical overview and open questions," *Proceedings of the IEEE*, pp. 1–27, 2024.

[22] F. Conceição, A. Pereira, M. Gomese, V. Silva, and R. Dinis, "Active intelligent surfaces for next generation radio systems: An overview on large intelligent surfaces and radio stripes," in *Proc. of 14th International Symposium on Communication Systems, Networks and Digital Signal Processing (CNSDSP)*, 2024, pp. 477–482.

[23] A. Fascista, B. J. B. Deutschemann, M. F. Keskin, T. Wilding, A. Coluccia, K. Witrisal, E. Leitinger, G. Seco-Granados, and H. Wymeersch, "Uplink joint positioning and synchronization in Cell-Free deployments with radio stripes," in *Proc. of IEEE International Conference on Communications Workshops (ICC Workshops)*, 2023, pp. 1330–1336.

[24] M. H. Moghaddam, S. R. Aghdam, N. Mazzali, and T. Eriksson, "Statistical modeling and analysis of power amplifier nonlinearities in communication systems," *IEEE Transactions on Communications*, vol. 70, no. 2, pp. 822–835, 2022.

[25] S. R. Aghdam, S. Jacobsson, U. Gustavsson, G. Durisi, C. Studer, and T. Eriksson, "Distortion-aware linear precoding for massive MIMO downlink systems with nonlinear power amplifiers," *arXiv preprint arXiv:2012.13337*, 2020.

[26] M. Majidi, A. Mohammadi, and A. Abdipour, "Analysis of the power amplifier nonlinearity on the power allocation in cognitive radio networks," *IEEE Transactions on Communications*, vol. 62, no. 2, pp. 467–477, 2014.

[27] D.-H. Shin and T.-K. Sung, "Comparisons of error characteristics between TOA and TDOA positioning," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 38, no. 1, pp. 307–311, 2002.

[28] H. Wymeersch, R. Amiri, and G. Seco-Granados, "Fundamental performance bounds for carrier phase positioning in cellular networks," in *Proc. GLOBECOM 2023 - 2023 IEEE Global Communications Conference*, 2023, pp. 7478–7483.

[29] S. Zhou, S. Wei, and J. Wang, "On-chip temperature compensation for small-signal gain variation reduction," *Micromachines*, vol. 13, no. 7, p. 1101, 2022.

[30] 3GPP, "Study on new radio access technology: Radio access architecture and interfaces," 3rd Generation Partnership Project (3GPP), Technical Report TR 38.901, 2017, release 17. [Online]. Available: https://www.etsi.org/deliver/etsi_tr/138900_138999/138901/17.01.00_60/tr_138901v170100.pdf

- [31] Y. Cao, M. Wojnowski, and B. K. Lau, "Circularly polarized Sub-THz antenna design for distributed deployment," in *Proc. 2024 18th European Conference on Antennas and Propagation (EuCAP)*, 2024, pp. 1–4.
- [32] Y. Cao, H. Aliakbari, P. Aghdam, and B. K. Lau, "Wideband single-layer D-band patch antenna with parasitic elements and superstrate loading," in *Proc. 2025 19th European Conference on Antennas and Propagation (EuCAP)*, 2025, pp. 1–5.



Erik G. Larsson (Fellow, IEEE) received the Ph.D. degree from Uppsala University, Uppsala, Sweden, in 2002. He is currently Professor of Communication Systems at Linköping University (LiU) in Linköping, Sweden. He was with the KTH Royal Institute of Technology in Stockholm, Sweden, the George Washington University, USA, the University of Florida, USA, and Ericsson Research, Sweden. His main professional interests are within wireless communications, signal processing, network science, and decentralized machine learning. He co-authored *Space-Time Block Coding for Wireless Communications* (Cambridge University Press, 2003) and *Fundamentals of Massive MIMO* (Cambridge University Press, 2016). He served as chair of the IEEE Signal Processing Society SPCOM technical committee (2015–2016), chair of the *IEEE Wireless Communications Letters* steering committee (2014–2015), member of the *IEEE Transactions on Wireless Communications* steering committee (2019–2022), General and Technical Chair of the Asilomar SSC conference (2015, 2012), technical co-chair of the IEEE Communication Theory Workshop (2019), and member of the IEEE Signal Processing Society Awards Board (2017–2019). He was Associate Editor for, among others, the *IEEE Transactions on Communications* (2010–2014), the *IEEE Transactions on Signal Processing* (2006–2010), and the *IEEE Signal Processing Magazine* (2018–2022). He received the IEEE Signal Processing Magazine Best Column Award twice, in 2012 and 2014, the IEEE ComSoc Stephen O. Rice Prize in Communications Theory in 2015, the IEEE ComSoc Leonard G. Abraham Prize in 2017, the IEEE ComSoc Best Tutorial Paper Award in 2018, the IEEE ComSoc Fred W. Ellersick Prize in 2019, and the IEEE SPS Donald G. Fink Overview Paper Award in 2023. He is a Fellow of EURASIP, a member of the Swedish Royal Academy of Sciences (KVA), and Highly Cited according to ISI Web of Science.



Dexin Kong (Graduate Student Member, IEEE) received his B.Eng. degree in communication engineering from Nanjing University of Posts and Telecommunications (NJUPT), Nanjing, China, in 2021, and M.Sc. degree in communication systems from Lund University, Lund, Sweden, in 2023. He is currently pursuing a Ph.D. degree with the Department of Electrical Engineering at Linköping University, Linköping, Sweden. His research interests include wireless communications, sub-terahertz communications, and radio-over-fiber technologies.



Diana Pamela Moya Osorio (M'16, SM'23) is currently Associate Professor at the Communication Systems Division, Department of Electrical Engineering, Linköping University, Sweden, and an EL-LIIT recruited faculty. Previously, she was Senior Research Fellow and Adjunct Professor at the Centre for Wireless Communications, University of Oulu, Finland. She received the B.Sc. degree in electronics and telecommunications engineering from the Armed Forces University, Ecuador, in 2008, and the M.Sc. and D.Sc. degrees in electrical engineering with emphasis on telecommunications and telematics from the University of Campinas, Brazil, in 2011 and 2015, respectively. From 2015 to 2022, she was an Assistant Professor with the Department of Electrical Engineering, Federal University of São Carlos, Brazil. From 2020 to 2023, she was also a Postdoctoral Researcher for the Academy of Finland. She has served as TPC and reviewer for several journals and conferences. Currently, she is Associate Editor of IEEE Communication Letters and IEEE Transactions on Information Forensics & Security. She also serves as working group leader for Trustworthy 6G at the Cost Action 6G-PHYSEC. Her research interests include wireless communications and sensing in general, signal processing for wireless communications and radar systems, physical layer security and trustworthy systems, and integrated sensing and communications.